Design and Control of Systems for Flexible Assembly

by

Ivan K. Oei

Submitted to the Department of Electrical Engineering and

Computer Science in Partial Fulfillment of the Requirements for the

Degree of

Master of Engineering in Electrical Engineering and Computer Science

at the

Massachusetts Institute of Technology

May 30, 1996

Author .....
Department of Electrical Engineering and Computer Science
May 30, 1996

Certified by .....
Professor Steven B. Leeb
Thesis Supervisor

Accepted by .....
Professor F.R. Morgenthaler
Chairman, Department Committee on Graduate Theses

Design and Control of Systems for Flexible Assembly

by

Ivan K. Oei

Submitted to the Department of Electrical Engineering and Computer Science on May 30, 1996, in partial fulfillment of the requirements for the Degree of Master of Engineering in Electrical and Computer Science

Abstract

As modern day products are becoming increasingly complex, the demands placed on the manufacturing facilities forces companies to seek alternative ways that are more capable yet less costly. Conventional methods of hand assembly or hard automation are becoming increasingly uneconomical especially in light of shorter product life cycles and diverse product families. Flexible assembly systems offer one solution. These systems would have the versatility to assemble a family of different products instead of just one type. However, such versatility comes with added complexity and problems with respect to process control and alignment.

This thesis addresses these problems associated with flexible assembly in the context of a system for bending products. An analysis of the system is conducted to derive a mathematical model of the system. Based on this model, a suitable controller is devised to optimize performance for speed and accuracy. This thesis also addresses two issues regarding alignment of components of a flexible assembly system. It describes the development of a secure clamping mechanism for two systems. It also documents the study of an optical sensing method for the purpose of fine part location.

Thesis Supervisor: Professor Steven Leeb
Title: Carl Richard Soderberg Assistant Professor of Power Engineering

Acknowledgments

The author would like to thank:

Marius and Suzy Oei for their unwavering love, support, and prayers, especially through the toughest and most precarious time in my life.

David Beal for being a good friend and the best partner I've ever had the pleasure of working with. My best wishes go out to you as you carry on the project; I look forward to the chance to work with you again.

Keith Drumheller and Tom Davis for their support and faith in my abilities.

Fred Cote for his assistance that came freely, patiently, and unpretentiously. You are truly a good person and a good resource.

Brian, Steve, Brandon, Wayne, and Jeremy - for their camaraderie and their Mech E advice. The times we had at the MI are ones that I will look back very fondly upon.

Leanna and Jessica for helping me out and for their smiles.

Jon, Brandy, Othon, Chung-Hsiu, Jessica, and Phu - friends in different places going different directions, but my good unforgotten friends nonetheless. Thank you.

Table of Contents

List of Figures

List of Tables

1 Introduction

This project involves the development of a flexible machine for the assembly of electronic products. The assembly process is usually a multi-step operation requiring a variety of different motions. Typically, products that are produced in small quantities are assembled by hand with the aid of various jigs and fixtures. This method of production is generally cost effective for production in low volumes, such as during the prototyping stages of product development or during the introduction of a product into the market. For large quantities, it is costly and slow. It may also lack a high degree of process repeatability since hand assembly of parts by workers is susceptible to errors.

For large quantities, companies will seek economies of scale by automating the entire process using high-speed, product-specific assembly machines. Automation generally provides a high rate of production, a low cost per unit, and a good amount of quality control. Unfortunately, the development of new machines or the adaptation of existing ones is a long and costly process. Furthermore, such machines are usually optimized for one product design, that is, to manufacture one part quickly and reliably. Thus, it is only cost-effective for products that are in high demand for long periods of time - or at least until the profits from sales recovers the costs of developing the machinery. For companies in volatile markets or with products that potentially become obsolete quickly, hard automation is a risky option.

Clearly, hand assembly and hard automation are not well suited for certain types of products. They are not suited for production of intermediate quantities where hand assembly becomes costly but not costly enough to justify the capital investment in expensive machinery. They are also not suited for products with high consumer demand but short lifecyles.

The basic idea behind flexible automation is to be able to assemble a variety of products with a minimum of fixed-tooling and product-specific parts and programming. A flexible assembly system must be capable of rapid product changeover times such that short run production of multiple types of parts are economically feasible. However, it must also be capable of performing medium length production runs reliably and do so at a reason-

able rate of output. It must be able to perform the variety of complex high-precision tasks that are necessary to assemble different products as well as tolerate some variation in the manufacturing environment and in the quality of the parts to be assembled.

However, such versatility also comes with added complexity and problems with respect to process control, reliability, and robustness. In order for a machine to perform multiple high-precision tasks, many components may be required to have multiple degrees of motion. Such flexibility requires a good understanding of the important system characteristics and non-trivial control efforts. Moreover, having to tolerate variability in the manufacturing environment and in the parts to be assembled requires that the system and the control scheme implemented be robust. In the course of developing a flexible assembly system, this thesis will examine the critical control issues and develop an analytical method of designing suitable controllers for such a system.

Efforts have been focused on the development of a control scheme for a module capable of performing multiple bends at different angles. In the assembly of products, it is often times necessary to bend the parts of a product to certain angles. Bending modules for product-specific systems are designed to be very good at performing bends for the intended product - perhaps forming accurate bends on multiple parts simultaneously, but they are generally incapable of performing bends for other products. A bender module for a flexible system, on the other hand, must be able to perform bends to almost any angle accurately although perhaps not as efficiently as hard automation. The mechanical design for this module was implemented at an earlier stage of this research endeavor [2]. Parameter characterization and the design of a suitable controller based on those parameters was left undone and is one focus of this thesis.

The bending module was meant to work together with another module that would be able to transport and present a product with unbent parts to the bender module. This transport module must be rigidly and accurately aligned with the bender module while it is working with it. Otherwise, the quality of the bends may be compromised as result of any misalignment that occurs. A mechanism for secure clamping between the transport module and the bender module was developed. The method for its design as well as test results of its clamping performance are documented here.

The last issue considered was one regarding the positioning of a product with respect to a module such as the bender or transport modules. As the demand for tighter geometric tolerances increases, the cost of making tools to build them also increases. This is because accurate location of a part requires accurate tooling. However, if the location of a part can be found by sensing for certain features and measuring them, then low tolerance tooling is unnecessary. Furthermore, if such a sensing method can be applied to a variety of parts, then the added flexibility makes it preferable to expensive, low-tolerance, fixed tooling. Some studies were conducted to measure the capability of a optical sensing method that made use of a common feature on the products.

The first four chapters of this thesis describes the derivation of a mathematical model for the bender module and the design of a suitable controller based on it. The fifth chapter documents the design for a clamping mechanism and the procedures whereby a similar mechanism can be created for a different application. The last chapter records the study of an optical sensing method for the purpose of part location.

Bender Module

An important function of the flexible assembly system is to bend parts to a specified angle. The bend may be as simple as a 90^{o} bend or as complicated as a smooth "s" shaped bend. Bending modules for product-specific systems are designed to be very good at performing bends for the intended product. They may even form accurate bends on multiple parts simultaneously. However, they are generally incapable of performing bends for other products. A bender module for a flexible system, on the other hand, must be able to perform bends to almost any angle accurately although perhaps not as efficiently a highly part-specific bender.

The performance of the bender module is defined by four specifications. The bender module must be able to bend to any angle between $0°$ and $90° \pm 0.5°$. It must bend the parts cleanly about a known center of rotation. The bender must also be able to compensate for the "springback" associated with different material thicknesses, compositions, hardness, and elasticity. Simply stated, springback is the characteristic of metal that is bent to revert back to its original shape. It will be studied in depth further in this thesis.

The unit must also bend to the desired angle without overshooting it. Unintentionally overbending the parts is undesirable since it would require bending the part in the opposite direction. Bending the metal back and forth to converge to the desired angle would tend to work harden it. Furthermore, since the bending process ought to be done as quickly and efficiently as possible, having to correct for an unintentional bend will clearly reduce the efficiency of the bender.

Finally, the bender module must be quick. Since it may be necessary for the module to bend each part on a multi-part product individually, making the system capable of running at high speeds is desirable to maximize the throughput of the machine. Of course, the above specifications of accurate bends and no overshoot must not be traded against operating speed.

The mechanical design for this module was implemented at an earlier stage of this research endeavor [2]. A mechanical arm capable of tightly gripping the parts has been designed and built. It has been coupled to a DC servo motor controlled by a computer

equipped with a servo control card. An automatic bending algorithm, however, had not been implemented.

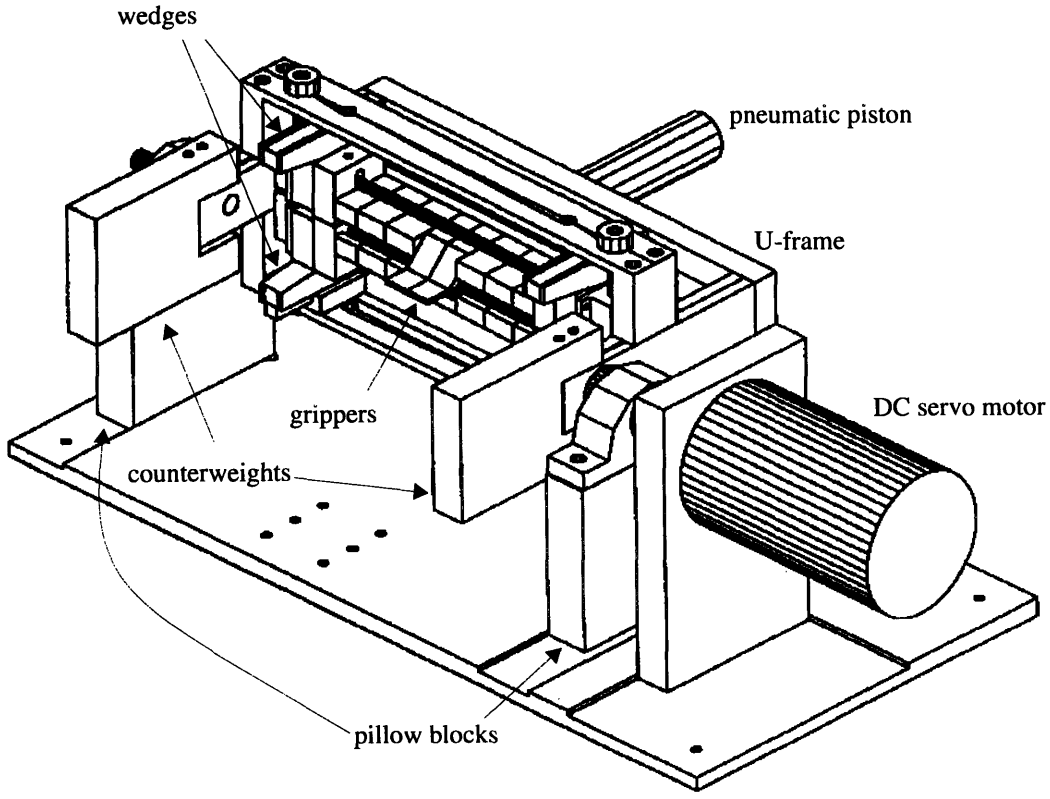
The objectives of chapters 2-4 are three-fold. First, create a theoretical model for the whole electromechanical system and verify it with experimental test results. Understanding the electrical and mechanical dynamics of the system is necessary so that an optimal compensator - optimal in the sense of meeting the three criteria above - can be created. Second, implement the appropriate compensator. Ultimately, implement an algorithm to perform bends automatically. Given a part (or parts) and a command specifying the desired bend angle, the machine should proceed to bend the parts given and compensate for the springback of the material.

2 Model of the mechanical system

2.1 Model of DC Servo

The mechanical system is essentially a DC servo motor that rotates an aluminum arm as shown in Figure 2.1.[1] The bending arm is an aluminum U-frame supported on either side by shafts in two pillow blocks, where one of the shafts is coupled to the motor. These shafts are secured in place by ball bearings, which rotate about the axis of the motor. A pair of grippers, designed to grip the parts uniformly and securely, are positioned at the center of the U-frame. These grippers are actuated by a pneumatic piston that is mounted to the bottom of the U-frame. Unfortunately, the placement of the piston shifts the center of gravity of the arm away from the axis of rotation. To move the center of gravity of the U-frame to the axis of rotation, counterweights were added to the sidebeams.

Figure 2.1: Bender module



1. Figure 2.1 as well as this system was prepared by Miguel Barrientos [2].

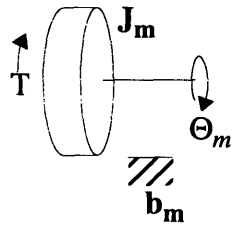


Figure 2.2: Free body diagram of DC servo

Assuming a current drive, the servo motor is modelled as a torque source that supplies a torque, **T**, through a rotor with mechanical parameters $\mathbf{J_m}$, the moment of inertia, and $\mathbf{b_m}$, a frictional damping coefficient. Utilizing Newton's Second Law of Motion (for rotational motion) the equation for the servo motor is

$$\mathbf{T} = \mathbf{J_m}\frac{d^2\Theta_m}{dt} + \mathbf{b_m}\frac{d\Theta_m}{dt}. \tag{2.1}$$

Thus, the torque supplied by the motor is equal to the product of the inertia with the rotational acceleration plus the product of the damping term with the rotational velocity. By taking the Laplace transform of this equation and rearranging variables, the transfer function relating angular shaft position to torque is

$$\frac{\Theta_m}{\mathbf{T}} = \frac{1}{s(\mathbf{J_m}s + \mathbf{b_m})}. \tag{2.2}$$

This system is represented schematically in Figure 2.2. It is a second order system with poles at s = 0 and **-b/J**.

2.2 Model of Bender Arm

The mechanical arm complicates the simple model for the servo motor. To evaluate its effects on the original servo motor model, it is useful to utilize a lumped parameter model of the arm. The rotating assembly can be thought of as a flywheel with inertia $\mathbf{J_l}$, and the pillow blocks as a single rotational damper with coefficient $\mathbf{b_l}$. Since the arm is very massive and not infinitely rigid, it will flex when large torques are applied. Thus, it is necessary to include a parameter that accounts for this springiness with some stiffness coefficient $\mathbf{k_c}$. Lastly, the parts bent by the module can be modeled as a torsional spring to

mechanical ground with constant stiffness $\mathbf{k_t}$. With these parameters, a torque balance equation can be written for the system

$$\mathbf{k_c}(\Theta_m - \Theta_l) = \mathbf{J_l}\frac{d^2}{dt}\Theta_l + \mathbf{b_l}\frac{d}{dt}\Theta_l + \mathbf{k_t}\Theta_l. \tag{2.3}$$

This equation relates the torque supplied by the motor to the product of the inertia of the load and its acceleration plus the losses seen by the load. The difference term in (2.3) between the motor shaft's position Θ_m and the load's position Θ_l reflects the material flexibility of the assembly. Rewriting (2.3) to group terms containing Θ_l together results in

$$\mathbf{k_c}\Theta_m = \mathbf{J_l}\frac{d^2}{dt}\Theta_l + \mathbf{b_l}\frac{d}{dt}\Theta_l + (\mathbf{k_t} + \mathbf{k_c})\Theta_l. \tag{2.4}$$

A frequency domain representation can be found by taking the Laplace transform of (2.4)

$$H_\Theta(s) = \frac{\Theta_l(s)}{\Theta_m(s)} = \frac{\mathbf{k_c}}{\mathbf{J_l}s^2 + \mathbf{b_l}s + (\mathbf{k_t} + \mathbf{k_c})}. \tag{2.5}$$

The usefulness of this equation will be made evident below.

2.3 Modified DC Servo Model

The equation for the servo motor (2.1) must also be modified. It must include the torsional spring term that couples the servo to the mechanical jig. The equation becomes

$$\mathbf{T} = \mathbf{J_m}\frac{d^2}{dt}\Theta_m + \mathbf{b_m}\frac{d}{dt}\Theta_m + \mathbf{k_c}(\Theta_m - \Theta_l) \tag{2.6}$$

where a free body diagram of it is shown in Figure 2.3.

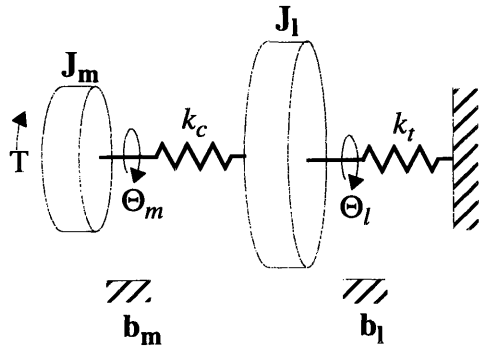


Figure 2.3: Free body diagram of modified DC servo model

Once again, it is useful to consider this equation in the frequency domain. Taking the Laplace transform of (2.6) gives

$$H_m(s) = \frac{\Theta_m(s)}{\mathbf{T}(s)} = \frac{1}{\left(\mathbf{J_m}s^2 + \mathbf{b_m}s + \mathbf{k_c}\right) - \mathbf{k_c}\left(\frac{\Theta_l(s)}{\Theta_m(s)}\right)}. \tag{2.7}$$

Substituting (2.5) into (2.7) gives a transfer function that is solely in terms of the motor and load mechanical parameters

$$H_m(s) = \frac{\mathbf{J_l}s^2 + \mathbf{b_l}s + (\mathbf{k_c} + \mathbf{k_t})}{\left(\mathbf{J_m}s^2 + \mathbf{b_m}s + \mathbf{k_c}\right)\left(\mathbf{J_l}s^2 + \mathbf{b_l}s + (\mathbf{k_c} + \mathbf{k_t})\right) - \mathbf{k_c}^2}. \tag{2.8}$$

This equation describes the transfer function between the torque input and the output angular position measured from the motor shaft. The system is fourth order. The zeroes of the function lie somewhere on the left half of the s-plane. Depending on the parameter values, $\mathbf{J_l}$, $\mathbf{b_l}$, $\mathbf{k_c}$, and $\mathbf{k_t}$, the zeroes could take on two negative real values or two complex conjugate values. The poles are also dependent on the parameter values, but they are less straightforward to determine.

If the arm was infinitely rigid, Θ_m equals Θ_l would be unity. Also, the motor and load inertias could be lumped as one inertia, $\mathbf{J_T}$. Similarly, the motor and load damping coefficients could be lumped as one coefficient, $\mathbf{b_T}$. The torque balance equation and transfer function $H'_m(s)$ that results would be second order of the form

$$\mathbf{T} = \mathbf{J_T}\frac{d^2\Theta_m}{dt} + \mathbf{b_T}\frac{d\Theta_m}{dt} + \mathbf{k_t}\Theta_m \tag{2.9}$$

$$H'_m(s) = \frac{1}{\mathbf{J_T}s^2 + \mathbf{b_T}s + \mathbf{k_t}}. \tag{2.10}$$

This equation is very similar to (2.2), the second order transfer function for the standard DC servo. In fact, this system would be second order; the poles of this function are at $s = \frac{-\mathbf{b_T} \pm \sqrt{\mathbf{b_T}^2 - 4\mathbf{J_T}\mathbf{k_t}}}{2\mathbf{J_T}}$. With the flexibility inherent in the structure, however, the system is essentially two second order systems coupled together.

For the fourth order system, it is difficult to determine the pole locations without knowledge of the values of the mechanical parameters. However, a sense of where the poles lie can be found by examining three cases:

1. $\mathbf{b_m}$, $\mathbf{b_l}$, and $\mathbf{k_t}$ are zero
2. $\mathbf{b_m}$ and $\mathbf{b_l}$ non-zero but less than $\mathbf{J_m}$ and $\mathbf{J_l}$, while $\mathbf{k_t}$ is zero
3. $\mathbf{b_m}$ and $\mathbf{b_l}$ non-zero but less than $\mathbf{J_m}$ and $\mathbf{J_l}$, and $\mathbf{k_t}$ small and non-zero.

This method was utilized in [16] to gain a qualitative understanding of a similar system.

The open loop characteristic polynomial of the system is

$$\mathbf{J_m J_l} s^4 + (\mathbf{J_m b_l} + \mathbf{J_l b_m}) s^3 + (\mathbf{b_m b_l} + \mathbf{J_m}(\mathbf{k_c} + \mathbf{k_t}) + \mathbf{J_l k_c}) s^2 + (\mathbf{b_m}(\mathbf{k_c} + \mathbf{k_t}) + \mathbf{b_l k_c}) s + \mathbf{k_c k_t}. \quad (2.11)$$

In the first case, this polynomial becomes

$$\mathbf{J_m J_l} s^4 + \mathbf{k_c}(\mathbf{J_m} + \mathbf{J_l}) s^2 \quad (2.12)$$

with two roots at the origin and two purely imaginary conjugate roots at $s = \pm j\sqrt{\frac{\mathbf{k_c}(\mathbf{J_m} + \mathbf{J_l})}{\mathbf{J_m J_l}}}$. Without any damping, the open loop system is marginally stable with all four poles on the $j\omega$-axis.

The second case, where $\mathbf{b_m}$ and $\mathbf{b_l}$ are non-zero but less than $\mathbf{J_m}$ and $\mathbf{J_l}$, is a more realistic situation. For the bender unit, the moments of inertia are quite large compared to the damping terms. This is particularly true for $\mathbf{J_l}$, the moment of inertia of the bender arm, since it is massive, as apparent in Figure 2.1, and loaded with lead counterweights for balance. Furthermore, the damping is kept to a minimum since ball bearings are used in the pillow blocks. The relative magnitudes of the inertias and damping coefficients in this case are very similar to those of the actual system. With the introduction of some small damping into the system, some of the poles move away from the $j\omega$-axis and into the left half s-plane. One would, however, expect them to be near the poles of the undamped system. The two imaginary poles, for example, will become complex conjugate poles as they move slightly off the $j\omega$-axis and into the left half plane. The characteristic polynomial for this case is

$$\mathbf{J_m J_l} s^4 + (\mathbf{J_m b_l} + \mathbf{J_l b_m}) s^3 + (\mathbf{b_m b_l} + \mathbf{J_m k_c} + \mathbf{J_l k_c}) s^2 + \mathbf{k_c}(\mathbf{b_m} + \mathbf{b_l}) s. \quad (2.13)$$

Based on (2.13), it is clear that one pole still resides at the origin. As for the last pole, since this model is for a real system, complex poles occur in conjugate pairs. Consequently, the remaining pole can only be a negative real pole.

In the final case, $\mathbf{k_t}$ is non-zero but small. This case accounts for the torsional spring constant that arises from bending parts. Since the parts are very malleable and small compared to the bender unit, it is reasonable to assume that $\mathbf{k_t}$ would also be small compared to the other parameters. The characteristic polynomial is (2.11). With $\mathbf{k_t}$ small, it has a minimal effect on the relative pole locations found in the second case. However, the most noticeable effect is that the pole that once was at the origin becomes a negative real pole. This is an important observation and has significant implications in the design of the compensator.

In summary, the transfer function that describes the bender module is

$$H_m(s) = \frac{\mathbf{J_l}s^2 + \mathbf{b_l}s + (\mathbf{k_c} + \mathbf{k_t})}{\left(\mathbf{J_m}s^2 + \mathbf{b_m}s + \mathbf{k_c}\right)\left(\mathbf{J_l}s^2 + \mathbf{b_l}s + (\mathbf{k_c} + \mathbf{k_t})\right) - \mathbf{k_c}^2}. \tag{2.14}$$

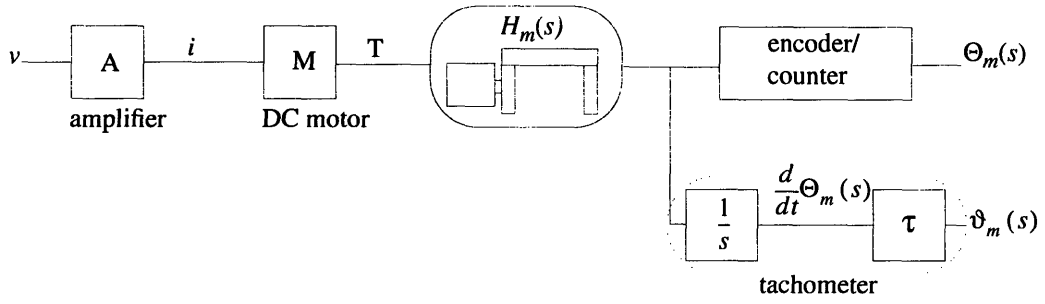
It is a fourth order system with two zeroes at $s = \frac{-\mathbf{b_l} \pm \sqrt{\mathbf{b_l}^2 - 4\mathbf{J_l}(\mathbf{k_c} + \mathbf{k_t})}}{2\mathbf{J_l}}$ and all poles in the left half s-plane. For $\mathbf{k_t} < \mathbf{b_l}, \mathbf{b_m} < \mathbf{J_l}, \mathbf{J_m}$, two of the poles will be complex conjugate numbers while the other two will take on real values.

2.4 Model with Additional Peripheral Components

The model is mostly complete. A few peripheral components must be included to account for the electrical input and output connections. The inputs to the system will be a voltage v, created by either a computer or a signal generator. This input is translated into a torque via a current amplifier and the DC servo motor as shown in Figure 2.4. The amplifier takes in the voltage and produces a current i that is proportional to the input voltage. Assuming that its bandwidth is much larger than the bandwidth of the mechanical system, this amplifier can be modelled as an element with a constant gain **A**.

The DC motor is an electromechanical component with mechanical and electrical characteristics that are relevant to the system model. As described in chapter 2.1, the motor is modelled by its moment of inertia $\mathbf{J_m}$ and damping coefficient $\mathbf{b_m}$. It can be

Figure 2.4: Model with peripheral components



viewed as an electromechanical element with constant gain M. It translates a current input i into torque T proportional to the current. Both characteristics are important and are modelled separately in Figure 2.4. The mechanical components are lumped with the system function $H_m(s)$ while the electrical component is modelled as a constant gain block.

The outputs of this system can be measured in two ways. An incremental encoder and a tachometer are both mounted to the motor shaft. The encoder produces a train of pulses that is proportional in number to the angular rotation that it observes. It thus discretizes a revolution of 360^o into a set number of pulses. A digital counter is used to count the pulses it sees. Given the number of pulses that corresponds to one revolution of the encoder, the amount of angular rotation can be measured. Together, the encoder and counter can be modelled as a quantizer that translates the amount of rotation of the motor shaft into an integer value. Since angular position control is an important requirement of this project, the encoder and counter are ideal for position measurement by a computer.

The tachometer provides an analog voltage that is proportional to the angular velocity. It can be modelled simply as an element with a "gain" of τ. In the block diagram of Figure 2.4, a differentiator block is included between the position output of the model of the bender unit Θ_m and the velocity input to the tachometer, $\frac{d\Theta_m}{dt}$. Since the time derivative in the time domain is multiplication by s in the complex frequency domain, the differentiator s is included to signify that the velocity input is the derivative of the position in the time domain.

Figure 2.5 is a block diagram of the bender unit with the peripheral electrical components. The transfer function *P(s)* that describes the relationship between the input voltage *V(s)* and output position $\Theta_m(s)$ is

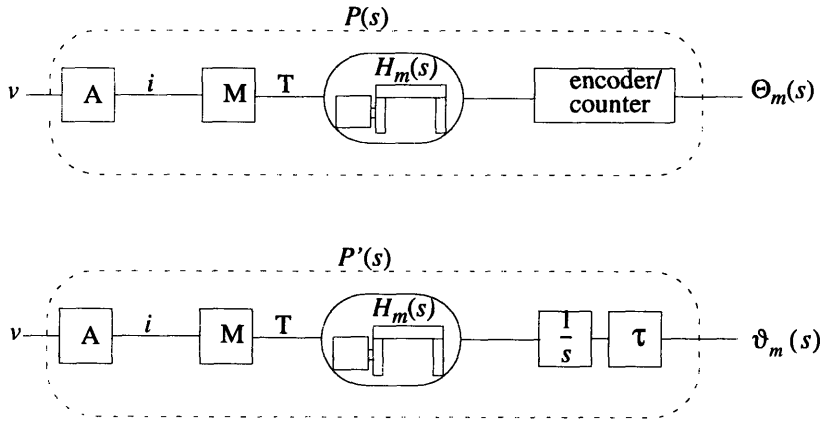
$$P(s) = \frac{\Theta_m(s)}{V(s)} = \frac{AM\left(J_l s^2 + b_l s + (k_c + k_t)\right)}{\left(J_m s^2 + b_m s + k_c\right)\left(J_l s^2 + b_l s + (k_c + k_t)\right) - k_c^2}, \tag{2.15}$$

where it is understood that the output has been quantized by the encoder. Similarly, the transfer function *P'(s)* describes the relationship between the input voltage and the velocity of the motor shaft measured by the tachometer $\vartheta_m(s)$

$$P'(s) = \frac{\vartheta_m(s)}{V(s)} = \frac{(AMs)\left(J_l s^2 + b_l s + (k_c + k_t)\right)}{\left(J_m s^2 + b_m s + k_c\right)\left(J_l s^2 + b_l s + (k_c + k_t)\right) - k_c^2}. \tag{2.16}$$

The model of the servo and the bender arm is complete. The choice of transfer function to use to describe the system depends on whether the tachometer will be used to measure velocity or the encoder to measure position, as shown in Figure 2.5.

Figure 2.5: Block diagram for *P(s)* and *P'(s)*



2.5 Experimental Verification

Determination of the pole and zero locations of the system can be done in a few different ways. Two were considered:

1. Measurement of the mechanical parameters in order to calculate the pole and zero locations, and

2. Measurement of the frequency response of the system to infer the pole and zero locations.

Calculation of the poles and zeroes using measurements of the mechanical parameters is generally possible for a system with a simple design. This method was attempted in [2]. For the bender unit, it is more difficult. Measuring the moment of inertia of the arm itself, for example, is not a simple task due to the odd geometry of the U-frame design. Furthermore, it is composed of parts made of different materials with different masses; the frame itself is aluminum, while the pneumatic piston is steel and the counterweights are lead. Consequently, the moments of each part of the arm must be calculated separately by measuring the mass, shape, and distance from the axis of rotation, and then summing them together. For some of the parameters, it is impossible to measure directly. The torsional spring constant $\mathbf{k_c}$, for example, is a function of the flexibility of the aluminum. It is also a function of the design since the arm can flex at the joints. At best, it can be approximated by considering the material properties of a rod of aluminum of roughly the same size. Finally, this method does not check the validity of the model derived above. It assumes that the model is correct since the poles and zeroes are found by finding the roots of the characteristic denominator polynomial and numerator polynomial, respectively, using the mechanical parameter values. For the above reasons, this first method was not attempted.

2.5.1 Measurement of Frequency Response

The second way of determining the pole and zero locations considered was by measuring the frequency response and inferring the locations from it. This was the method that was attempted and is essentially simple. Using a range of frequencies that roughly span the dynamic range of the mechanical system, drive the system with sinusoidal inputs and measure the relative magnitude and phase of the outputs. Any sharp spikes or drops in the magnitude response indicate the existence of complex conjugate poles and zeroes, respectively. For real poles and zeroes, the general slope of the magnitude response and the phase at different frequencies provide good indicators as to where the real poles and zeroes are [14].

For this experiment, a frequency generator was used to provide sinusoidal input waveforms of frequencies between 4.5 radians/second (0.72 Hz) and 2586 radians/second (412

Hz). Frequencies below 4.5 radians/second were not tried since at such low frequencies, the bender arm slammed against the mechanical limits of motion providing little useful information. Frequencies above 2586 radians/second were not tried because at such high frequencies, the output became too small to discern from the noise in the signal. This voltage signal was brought into the current amplifier to drive the motor.

The actual input signal monitored on an oscilloscope, however, was the current into the motor. This was done by attaching a current probe to one of the input lines to the motor. Measuring the current was done as a precaution to ensure that any unmodelled dynamics of the amplifier did not affect the experiment or at least were made evident upon observation of the current to motor.

The output monitored on the oscilloscope was the velocity of the motor shaft taken from the tachometer mounted on it. Consequently, the transfer function that was relevant to this experiment was *P'(s)* without the amplifier gain **A**, *P''(s)*

$$P''(s) = \frac{\tau \mathbf{M} s\left(\mathbf{J}_l s^2 + \mathbf{b}_l s + (\mathbf{k}_c + \mathbf{k}_t)\right)}{\left(\mathbf{J}_m s^2 + \mathbf{b}_m s + \mathbf{k}_c\right)\left(\mathbf{J}_l s^2 + \mathbf{b}_l s + (\mathbf{k}_c + \mathbf{k}_t)\right) - \mathbf{k}_c^2}. \tag{2.17}$$

The tachometer was chosen for this experiment over the encoder because the tachometer provided an analog measurement that could be easily observed on the oscilloscope and with confidence, known to be directly related to the velocity. In using the encoder and counter available, any time-related measurements of the output would have been susceptible to errors due to interrupt latencies and other delays associated with the operation of a computer. The magnitude of the output measurements would have also been susceptible to small quantization errors as a result of discretizing the measurement of the position. Based on the qualitative analysis in section 2.3 and on the above function, the experiment should reveal four poles: two real and two complex conjugate. It should also reveal two zeroes from the mechanical system function $H_m(s)$ and one from the differentiator at the origin. The frequency of the input, the magnitudes of the input current and output velocity, and the time delay between the two signals were measured for a range inputs at various frequencies. The details of this experiment are described in the "Flexible Assembly Project Software Technical Report" [11].

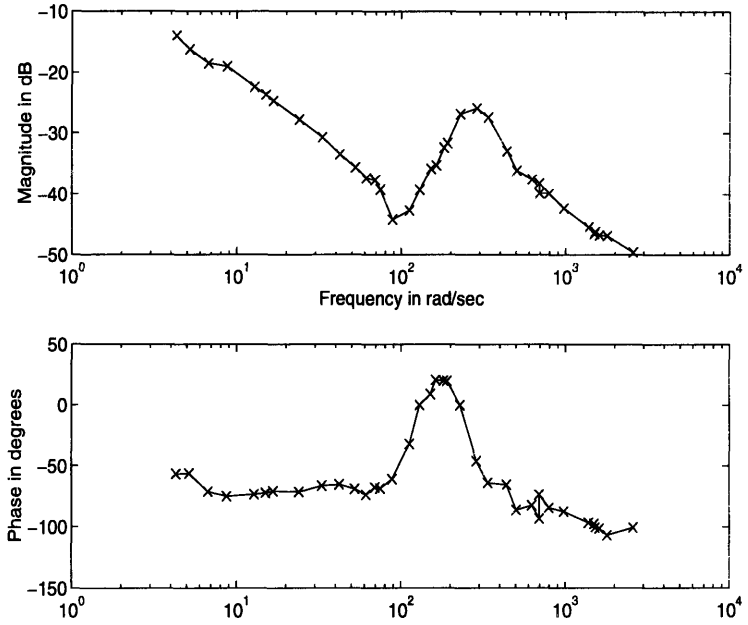


Figure 2.6: Measured frequency response of function $P''(s)$

The results of the experiment are shown in Figure 2.6. The ratio of the amplitude of the current waveform to the amplitude of the velocity was plotted in decibels versus the input frequency as the "Magnitude." The time difference between the two waveforms was normalized by the period of the input, scaled 180^{o}, and plotted versus the input frequency as the "Phase."

The drop in the magnitude around 100 radians/second indicates the presence of two complex conjugate zeroes, while the spike around 260 radians/second indicates the presence of two complex conjugate poles. Furthermore, the slope of the magnitude is roughly -20dB/decade of frequency, excluding the regions around the complex conjugate pole and zero, and the phase is around -90^{o}. This suggests that system has a pole outside the frequency range of the experiment.

After some curve-fitting attempts on Matlab using $P''(s)$ as the base function, it was found that the best fit to the data produces an estimate, $\hat{P}''(s)$, in factored form of

$$\hat{P}''(s) = \frac{7\,(s+30+100j)\,(s+30-100j)}{(s+3)\,(s+55+260j)\,(s+55-260j)}. \tag{2.18}$$

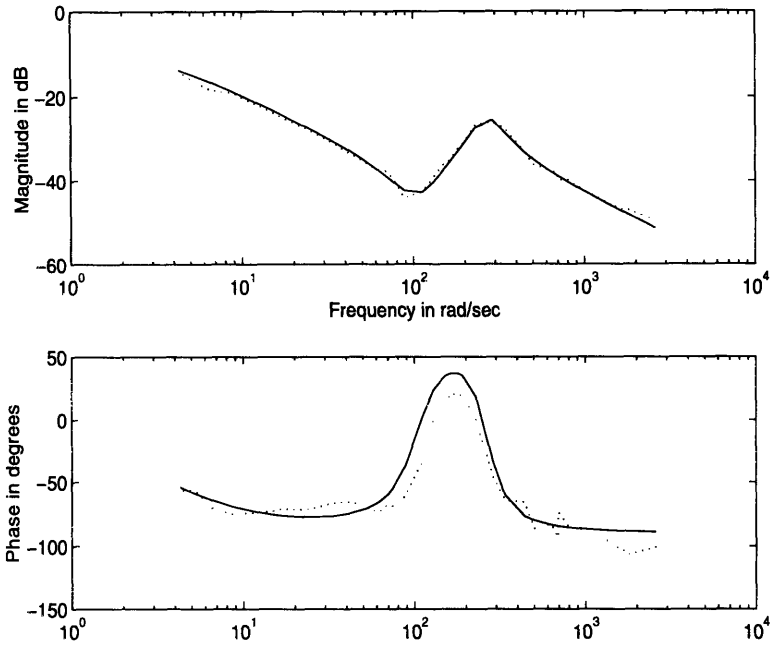


Figure 2.7: Frequency response of $P''(s)$ and $\overline{P}''(s)$

The estimated response $\hat{P}''(s)$ (solid line) is overlaid on the measured response $P''(s)$ (dotted line).

Thus far, the empirical data seems to support the form of the theoretical transfer function $P''(s)$, which is plotted in Figure 2.7. The two zeroes have been found and determined to be a complex conjugate pair at $s = 30 \pm 100j$. The two complex conjugate poles at $s = 55 \pm 260j$ are also evident as well as the existence of a real pole at $s = 3$. The open loop gain was also determined to be 7. The last mystery left is the location of the last pole; the zero is not a concern since it is an artifact of the tachometer and not inherent in the transfer function for the servo and mechanical arm.

2.5.2 Measurement of Steady State Error

Given all of the other poles and zeroes have been found, the last pole can be found by examining the steady state error due to a step input to the closed-loop system under unity feedback. The model of the servo and arm described by the function, $P(s)$, is considered

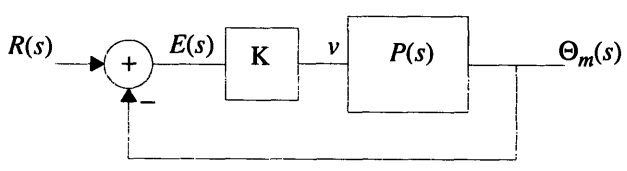


Figure 2.8: System under unity feedback

$E(s)$, the output of the summing junction, and proportional gain K are labelled.

"under closed-loop control with unity feedback" when its position output, $\Theta_m(s)$, is compared with a position command input, say $R(s)$, and fed into the input of $P(s)$. This configuration is shown in Figure 2.8. The transfer function relating the output of the summing junction, $E(s)$, to the input, $R(s)$, is of particular interest since it describes the "error", the difference between the output position and the input position. Using $P(s)$, the values of the poles and zeroes found above, and assuming that a proportional gain K is used, the error function is

$$\frac{E(s)}{R(s)} = \frac{1}{1+KP(s)} = \frac{(s+\alpha)\ (s+3)\ (s+55+260j)\ (s+55-260j)}{(s+\alpha)\ (s+3)\ (s+55+260j)\ (s+55-260j) + 7K\,(s+30+100j)\ (s+30-100)} \quad . \quad (2.19)$$

The steady state error of this closed-loop system to a step input is the difference between the output position and the input position long after the step has occurred, i.e., as time goes to infinity. That is, the steady state error ε can be expressed as

$$\varepsilon = \lim_{t\to\infty}\left(\frac{e(t)}{r(t)} * u(t)\right) \quad (2.20)$$

where $\frac{e(t)}{r(t)}$ is the time domain representation of (2.19), "*" denotes convolution, and $u(t)$ is the unit step function. By the Final Value Theorem [13], this limit in the time domain can be equivalently expressed as

$$\varepsilon = \lim_{s\to 0} s\left(\frac{E(s)}{R(s)}U(s)\right) = \lim_{s\to 0}\left(\frac{E(s)}{R(s)}\right) \quad . \quad (2.21)$$

where $U(s) = \frac{1}{s}$ is the Laplace transform of $u(t)$. The steady state error for the closed loop system given a proportional gain K can be found experimentally. By substituting (2.19) into (2.21) and evaluating the limit, (2.21) becomes

$$\varepsilon = \frac{\alpha(3)\ (55+260j)\ (55-260j)}{\alpha(3)\ (55+260j)\ (55-260j) + K7\,(30+100j)\ (30-100)} = \frac{211875+\alpha}{211875\alpha + 76300K} \quad . \quad (2.22)$$

From this equation, the location of the pole is

$$\alpha = \frac{76300K\varepsilon}{211875\,(1-\varepsilon)} \quad . \quad (2.23)$$

If the error ε and the gain K are known, then the location of the last pole can be found.

For this experiment, using the encoder was acceptable since the measurement of interest was not related to any period of the input or any time delays. Using a proportional gain $K = 4$, the steady state error to a unit step was $\varepsilon = 0.1453$ which gives $\alpha = 0.245$ radians/second. Thus the complete mathematical model of the system is

$$\hat{P}(s) = \frac{7\,(s+30+100j)\,(s+30-100j)}{(s+0.245)\,(s+3)\,(s+55+260j)\,(s+55-260j)}. \tag{2.24}$$

The Bode plot of this function is shown in Figure 2.9 From this point on, for simplicity of notation, the estimated transfer function will be noted as $P(s)$.

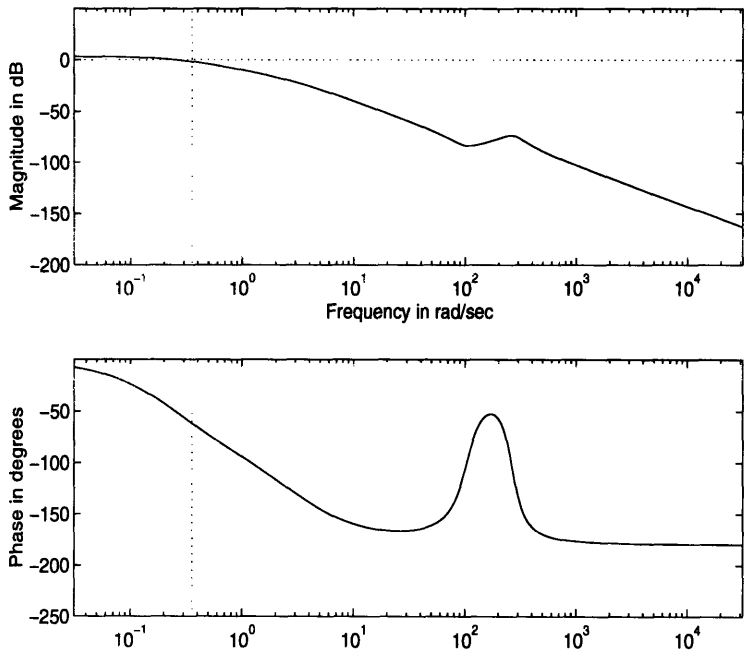


Figure 2.9: Complete estimated model of the bender module, $\hat{P}(s)$

3 Compensator Design in Continuous Time

With a model of the plant, a compensator can be created so that the response of the whole system met the requirements for accuracy, no overbending, and low bending times. This requires that the system have a fast step response to the desired position with no overshoot and zero steady-state error. Although the final compensator will be implemented in discrete time, a continuous-time compensator is first designed in order to understand the necessary features of a compensator in discrete time.

First, it is important to know what the step response of the closed-loop system by itself is in order to know what an appropriate compensator must do for the system to meet the specifications. By closing the loop with unity feedback, the (estimated) transfer function $P(s)$ of the system has the form which follows Black's Formula [14]:

$$P_{cl}(s) = \frac{P(s)}{1 + P(s)} \tag{3.1}$$

The step response of this transfer function, as shown in Figure 3.1, is extremely slow. The

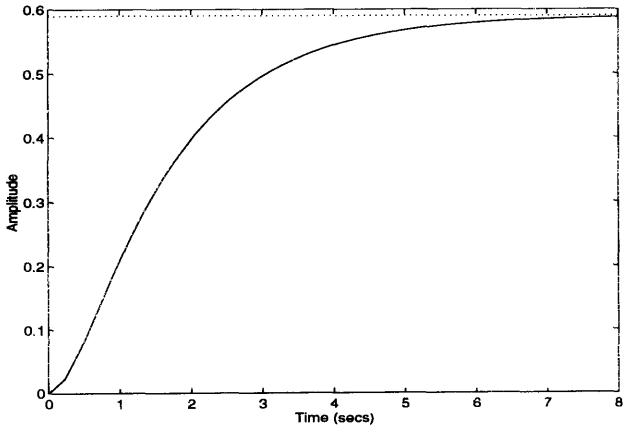


Figure 3.1: Step response of system under unity feedback

rise time, measured as the time from 10% to 90% of the final value, is 6 seconds - far below the minimum 1 second rise time that is specified. Furthermore, it has a very large steady-state error. However, the system is well-damped, which is a desirable attribute to maintain.

The steady-state error is due to the fact that the open loop system does not have infinite DC gain. Furthermore, upon examination of the Bode plots of the plant, as shown in Fig-

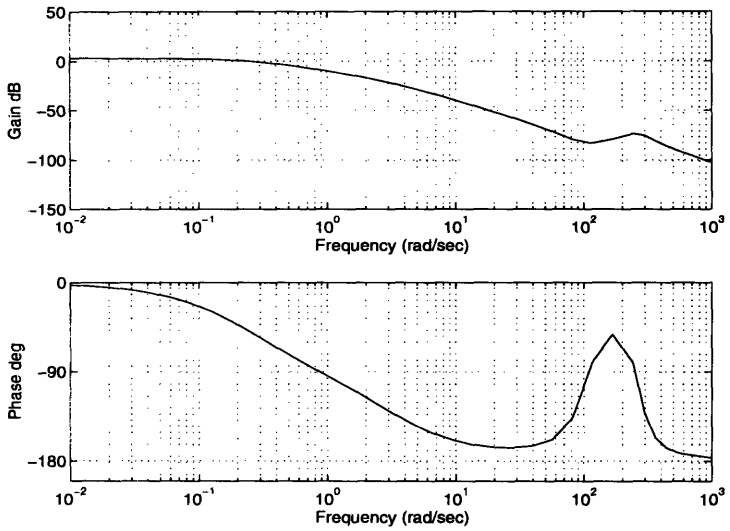


Figure 3.2: Bode Plot of $P(s)$

ure 3.2, it is clear that the slow rise time and well-damped response is due to the low bandwidth of the system and large phase margin. For a system whose Bode magnitude plot generally decreases, the 3dB bandwidth and the crossover frequency are related measures of responsiveness of the system. Increasing one generally increases the other. From the Bode plots, the crossover frequency is at 0.37 radians/second and the phase margin is 83^{o}. Since the phase margin is so large, it is possible to shift the crossover frequency by increasing the gain of the system and not sacrifice the overdamped response. To maintain a well-damped response, a conservative lower limit of the phase margin is roughly 60^{o}. The large peak in the phase response around 80 and 300 radians/second may seem alarming, but since it occurs far below crossover, the effect on the step response is negligible. At the frequencies where the peak occurs, any input signal at those frequencies is significantly attenuated. As long as the magnitude at those frequencies stay far below unity, the spike in the phase will not affect the step response. Based on these insights any compensator that helps the overall system meet the performance specifications must:

- remove the steady-state error in the step response
- increase the crossover frequency in order to increase the bandwidth of the system
- maintain a phase margin above 60^{o} so that the step response will not overshoot
- keep the magnitude less than unity between 80 and 300 radians/second.

3.0.1 Proportional Gain Compensation

Using a simple proportional gain compensator is a good first attempt at improving the performance of the system. However, it cannot improve the phase margin *and* increase the

bandwidth simultaneously. It also does little to reduce the steady-state error when using relatively low gains. This compensation scheme cannot directly change the phase response of the system; it merely shifts the magnitude response vertically based on the gain used. It can improve the phase margin indirectly only when the crossover frequency is decreased. Its effect on the steady-state error is only significant as the gain approaches very large values. That is, for the infinite (or at least very large) DC gains that are necessary in order to have zero steady-state error, the proportional gain compensator must have a very large gain. However, this is not possible if a 60^{o} phase margin is to be preserved. The gain that increases the crossover frequency while keeping the phase margin above 60^{o} is 2.5. The step response is shown in Figure 3.3. The rise time is decreased to 3 seconds, and the

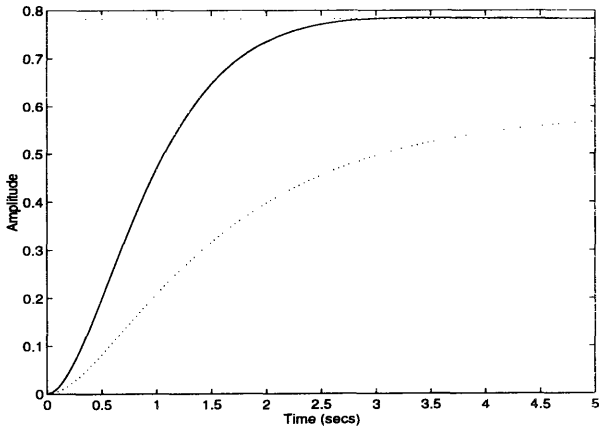


Figure 3.3: Step response of closed-loop system with proportional gain of 2.5

The step response with a gain of 2.5 (solid line) is compared with the step response with unity-gain (dotted line)

crossover increased to 0.85 rad/sec. This rise time is still below the specification and the steady-state error is quite large.

3.0.2 Lead Compensation

Improving the phase margin of the system while increasing the crossover frequency stipulates that the phase response must be changed at least around where the new crossover frequency is desired. Reducing the steady-state error of the system also implies that the magnitude response must be changed around the origin or very low frequencies. This cannot be done by merely changing the gain of the system; rather, it requires that dynamic components (poles and zeroes) be added to the system. One way of "cheating" the relationship between the magnitude and phase responses around crossover is to introduce a zero. The positive phase contribution of the zero, appropriately placed, would serve to

improve the phase margin of the system; this is known as lead compensation. To give the system infinite DC gain, a pole must be added at the origin, along with a zero to prevent the system from becoming unstable; this is known as lag compensation [7]. Since both compensators affect different sections of the frequency response, they can be designed independently as long as when they are integrated with the system, the appropriate gain term is used to preserve the desired crossover frequency.

The lead compensator introduces a zero into the system to improve the phase margin. Since a zero with a magnitude that approaches infinity as the angular frequency ω approaches infinity has practical problems, a more useful form of a lead transfer function is

$$K_{lead}(s) = K_d \cdot \frac{\alpha\tau_d s + 1}{\tau_d s + 1} \tag{3.2}$$

That is, the compensator introduces a zero at $s = \frac{1}{\alpha\tau}$ and a pole at $s = \frac{1}{\tau}$. The constants α and τ, are chosen such that the phase contribution of the zero is effective at the region of interest. The role of the pole is to undo what the zero did and thus localize the compensator's effect to a small region of the frequency response. To increase the crossover frequency to 4 radians/second, a proportional gain of 20 is required. The phase margin of the plant with just a proportional gain compensator drops to 37°. A lead compensator must add roughly 30° of phase to bring the phase margin above 60°. To increase this phase

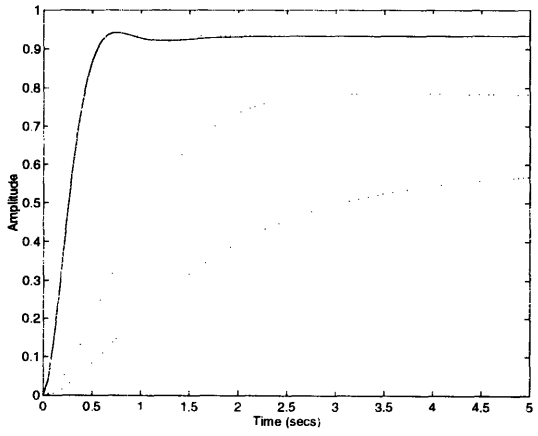


Figure 3.4: Step response of closed-loop system under lead compensation

The response with lead compensation (solid line) is compared with previously determined step responses (dotted lines).

while maintaining the current crossover frequency, $\alpha = 3$, $\tau_d = 0.0421$, and $K_p K_d = 11.54$ is required. Consequently, the compensator has the form

$$K(s) = \frac{9.8\,(0.1402s + 1)}{0.0421s + 1} \quad . \tag{3.3}$$

The step response of the lead-compensated system, shown in Figure 3.4, has no overshoot and reaches steady state (+/- 5%) immediately. The rise time of the system is 0.5 seconds, which exceeds the original specifications. Figure 3.5 also shows how the compensated and uncompensated system Bode plots differ. The phase margin has been improved clearly around crossover. In fact, the biggest improvement of the phase is around crossover, as desired. However, the 6% steady state error between the actual and desired values in the response still needs to be addressed. In terms of desired and actual bend angles that the bender unit would make, for a 90° bend, the actual bend angle would be 84.6° - far below the ±0.5° tolerance allowed about the desired bend angle.

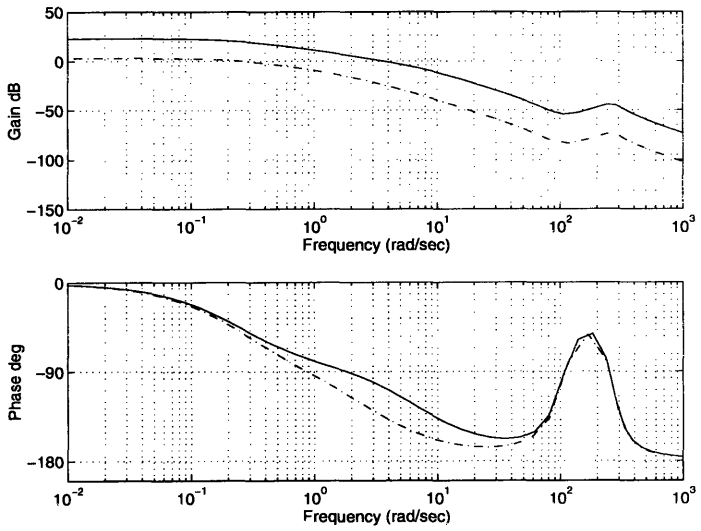


Figure 3.5: Bode plot of system with lead compensator

Gain and phase plots of lead compensated system (solid line) are compared with that of the original system (dotted line). Crossover is around 4 radians/second with a phase margin of 67°.

3.0.3 Lead-lag compensation

A lag compensator is typically used to increase the open loop gain at low frequencies while maintaining the current phase margin [7]. It has the form

$$K_{lag}(s) = K_i \cdot \frac{\tau s + 1}{\alpha \tau s + 1} \quad . \tag{3.4}$$

For infinite DC gain, the pole must be placed at the origin, so (3.4) takes the form

$$K_{lag}(s) = K_i \cdot \frac{\tau_i s + 1}{s} \tag{3.5}$$

and is more commonly known as an integral compensator for the existence of the integrator $\frac{1}{s}$. The purpose of the zero at $s = \frac{1}{\tau}$ is similar to that of the pole in the lead compensator: to localize the effect of the compensator to a region of the frequency response. To preserve the phase margin and crossover frequency that was achieved by using the lead compensator, $K_i = 0.89$ and $\tau = 3$ radians/second is required. The total compensator takes the form:

$$\begin{aligned} K_{total}(s) &= K_d K_i \cdot \frac{\alpha\tau_d s + 1}{\tau_d s + 1} \cdot \frac{\tau_i s + 1}{s} \\ &= \frac{8.722\,(0.1402s + 1)\,(3s + 1)}{(0.0421s + 1)\,s} \end{aligned} \tag{3.6}$$

The resulting step response is shown in Figure 3.6 below. The lead-lag compensator allows the performance specifications to be met. The rise time is roughly 0.5 seconds, the steady state error is 0, and there is no overshoot in the step response. The Bode plots in Figure 3.7 of the lead-lag compensated system shows that magnitude decreasing from low frequencies. The crossover frequency is around 4 radians/second with a phase margin of 67^o.

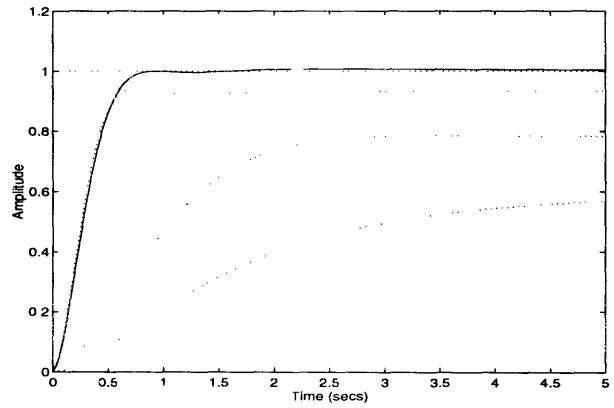


Figure 3.6: Step response of closed-loop system under lead compensation

Response under lead-lag compensation (solid line) is compared to responses previously found (dotted line).

Steady state error is 0.

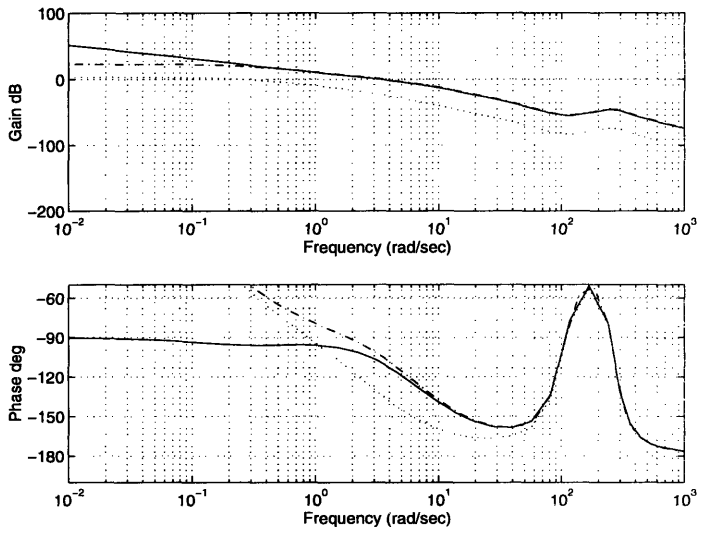


Figure 3.7: Bode plot of system with lead-lag compensation.

Although the phase response (solid line) is limited to -90° at low frequencies, the magnitude (solid line) is increasing at lower and lower frequencies. Previously determined responses (dotted line) are plotted for comparison.

Further attempts were made to improve the step response of the system, but they were unsuccessful. One way of improving the response is by increasing the bandwidth, that is, make it faster by using a larger gain which increases the crossover frequency. However, merely increasing the gain K_d of the current lead compensator reduces the phase margin of the overall system. A compensator that provides more phase can be used to keep the phase margin above 60° for K_d that increases crossover. In order to increase the phase contribution of the compensator, the zero must be brought closer to the jω-axis and thus, closer to the pole at the origin. But this introduces a new problem. The pole from the origin has moved along its locus into the left half s-plane toward the zero of the compensator. Since the pole and zero are in close proximity of one another, they act as a doublet. This results in the long tail toward the final value. Figure 3.8 shows this problem for $\alpha = 9$, $\tau = 0.08094$, and $K_d = 0.3333$. Crossover frequency has been increased to 10 rad/sec and the

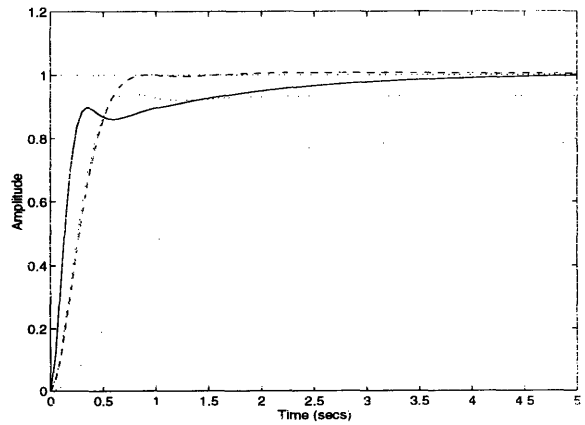


Figure 3.8: Step response of lead-lag compensated system with phase margin of 90°.

The slow rise to the final position of this system (solid line) is very undesirable.

phase margin up to 90^{o}. However, the step response depicts a slow creep to the steady state value.

From this derivation of a compensator in the continuous-time domain, it is clear that the appropriate compensator for the discrete-time implementation is a lead-lag compensator. The lead compensator is useful for its positive phase contribution since additional phase around the crossover frequency is required to improve the phase margin. On the other hand, the lag compensator is useful for the additional gain it provides to the frequency response at low frequencies since infinite DC gain is required for a zero steady state response. For both compensators, they must be designed to influence only the region of the frequency response they are to improve - for the lead compensator, around crossover and for the lag, at DC. As a beneficial consequence, they can be designed independently. It is also clear that if the zero of the lead compensator is not placed appropriately, it will act in conjunction with the pole as a doublet and affect the step response adversely. Thus, the lead compensator must be designed to provide the phase necessary to improve the phase margin but not create an effective pole-zero doublet.

3.1 Continuous-time to Discrete-time Transform

The compensator for this system will be implemented in discrete-time by a computer. A discrete controller uses sampled values of the output position of the machine to generate a time-evolving impulse train as the command signal. An Analog-to-Digital converter (ADC) performs the sampling of the output, and a Digital-to-Analog Converter (DAC) translates the digital impulse train to an analog signal that is used to control the module. It is necessary to include any effects that the ADC and DAC have on the system.

For this system, the angular position is measured digitally by an incremental encoder that is coupled to the back end of the motor shaft. The computer calculates the position by counting the number of pulses it sees from the rotation of encoder. Its measurement is based on predefining a point as the origin, or "home" position, and on knowing what the spacing between pulses is. The encoder provides two sets of pulses, one that is 90^{o} out of phase with the other. By comparing the phase lag or lead of one signal versus the other, the computer can determine the direction of rotation of the motor.

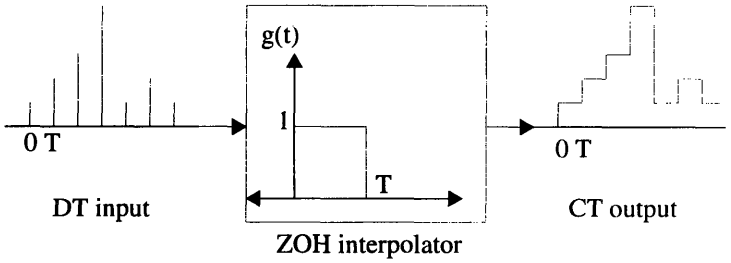


Figure 3.9: Zero-order hold interpolator

For this bender unit, a DAC is used to translate a 16-bit number to some voltage within a predefined range and convert them from an impulse train to a quantized continuous waveform. The DAC maintains its output voltage at the level commanded by the most recent impulse value until the next value arrives. This digital-to-analog reconstruction produces a "staircase" waveform and is known as zero-order hold (ZOH) interpolation. Better interpolators exist, such as a first-order hold interpolator where the output of the DAC is ramped up or down between impulse values - based on the rate of change of the magnitude between an impulse and the following impulse. But they are less common. They are also unnecessary for this system, especially when the effects of the zero-order hold interpolator can be included in the system model. In fact, it is necessary to account for any reconstruction interpolator in the discrete-time model that is used unless the reconstruction from digital to analog was perfect.

The ZOH interpolator can be considered in the continuous-time domain as a filter whose impulse response is a unity-magnitude pulse with a duration equal to the sampling period of the DAC, as shown in Figure 3.9. Thus its transfer function, $G_{zoh}(s)$, is.

$$G_{zoh}(s) = \frac{(1 - e^{-sT})}{s}$$

where T is the sampling period of the DAC.

Using $P(s)$ to represent the transfer function of the "plant" - motor and arm - and $x_o(s)$ and $x_i(s)$ to represent the output and input to the ZOH filter cascaded with the plant, the result is

$$\begin{aligned} x_O(s) &= G_{zoh}(s) \cdot P(s) \cdot x_i(s) \\ &= \frac{(1-e^{-sT})}{s} \cdot P(s) \cdot x_i(s) \end{aligned}$$

Thus the transfer function, $P_{zoh}(s)$, that describes the plant and ZOH filter is

$$P_{zoh}(s) = \frac{x_o(s)}{x_i(s)} = (1-e^{-sT}) \cdot \frac{P(s)}{s}. \tag{3.7}$$

This is the continuous time model that the compensator sees through the ZOH interpolator. This equation can be interpreted as the delayed step response of $P(s)$ subtracted from the undelayed step response. Since a discrete time transfer function is what is desired, the Z-transform can be found

$$P_{zoh}(z) = (1-z^{-1}) \cdot Z\{L^{-1}\{\frac{P(s)}{s}\}\}, \tag{3.8}$$

where $Z\{\}$ and $L^{-1}\{\}$ represents the z-transform and inverse Laplace transform of the terms in the brackets, respectively. (1.27) states that from the transform of the step response of $P(s)$, the z-transform of the DT step response can be found. This z-transform is then multiplied by $(1-z^{-1})$, the discrete-time equivalent of the $(1-e^{-sT})$ term in (1.26). Note that it is also the reciprocal of the transform for a discrete-time unit step. In effect, $P_{zoh}(z)$ is the z-transform of the DT step response of $P(s)$, normalized by the z-transform for a DT unit step [6].

The above procedure also embodies the procedure for a *step-invariant transform*, described in detail in [1]. From continuous-time to discrete-time, the step response is preserved across the transformation.

With a sampling period of $\mathbf{T} = 10^{-4}$ seconds, the continuous-time model $\hat{P}(s)$ becomes

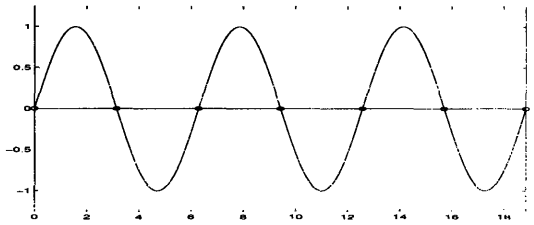


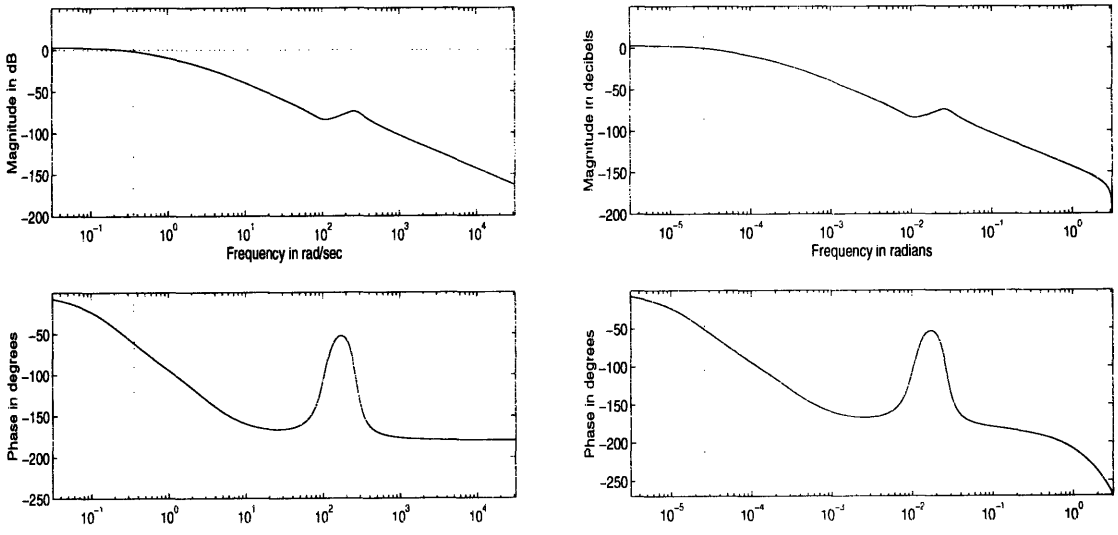
Figure 3.10: Zero due to sampling

$$\hat{P}_{zoh}[z] = \frac{(3.4936\times10^{-8})z^{-1} - (3.4785\times10^{-8})z^{-2} - (3.4809\times10^{-8})z^{-3} + (3.4666\times10^{-8})z^{-4}}{z^{-1} - 3.9880z^{-2} + 5.9648z^{-3} - 3.9655z^{-4} + 0.98874z^{-5}} \quad (3.9)$$

after applying the step-invariant transform. The poles in the z-plane are at $z = 0.9997, 0.99997$ and $0.99418 \pm 2.5854\times10^{-2}j$. The zeroes are at z = -1 and $0.90695 \pm 9.9698\times10^{-3}j$. The zero at z = -1 (or equivalently at π radians) is the result of sampling. Input signals with a period equal to integer multiples of the sampling period are not seen after discrete time sampler. Figure 3.10 depicts how a continuous time sinusoidal input with a period of twice the sampling period is "lost" by the discrete time sampler.

In comparing the frequency responses of the CT system and the DT system in Figure 3.11, it is clear that they are equivalent. The horizontal axes are different by a factor of 10^{-4} because in continuous-time frequency is measured in radians per second, whereas in discrete time frequency is measured in radians. The factor of 10^{-4} is attributable to the sampling frequency.

Figure 3.11: Continuous and discrete time Bode plots. The CT response (left) is equivalent to the DT response (right).



4 Compensator Design in Discrete Time

4.1 CT to DT Design by the Bilinear Transform

With a discrete-time model of the plant, a compensator in discrete-time can be created. In the previous section, it was shown that a continuous-time compensator exists that enables the bender module to perform bends within the specifications for its performance. This section will show that the discrete-time version of this lead compensator, if it can be implemented, is what is desired.

Unlike the plant which must be translated through the use of the step-invariant transform, the compensator has to be translated through an impulse invariant transform. There is no ZOH interpolator to consider. Rather, since the compensator will be implemented directly as a discrete-time controller, the emphasis is to preserve the pole-zero relationship across the two domains. The bilinear transform accomplishes this by non-linearly mapping the entire jω-axis in the s-plane to one revolution of the unit circle in the z-plane [12]. It corresponds to replacing s by

$$s = \frac{2}{T} \cdot \frac{1 - z^{-1}}{1 + z^{-1}}$$

that is,

$$H[z] = H\left(\frac{2}{T} \cdot \frac{1 - z^{-1}}{1 + z^{-1}}\right) \tag{4.1}$$

where $H[\ ]$ and $H(\ )$ denotes the discrete-time and continuous-time transfer functions, respectively. This mapping does warp the frequency response (both the magnitude and phase) of the controller since the continuous response over all frequencies are mapped to Ω between 0 and π radians, i.e. between 0 and $\frac{1}{2T}$ radians/second where T is the sampling period. However, the placement of the pole and zero relative to one another and the unit circle is what must be preserved, not the absolute placement in radians. Consequently, the warping is acceptable, and the bilinear transform is the best method for the emulation of the CT controller by a DT controller [12]. Through application of this transform, (4.1), to (3.6) for a sampling period of T = 10^{-2} seconds and simplifying the result, the discrete-time transfer function from the continuous-time function becomes

$$K_{total}[z] = K_{lead}[z]\,K_{lag}[z] \tag{4.2}$$

$$= 25.6\left(\frac{1-\alpha z^{-1}}{1-\beta z^{-1}}\right)\left(\frac{1-\gamma z^{-1}}{1-z^{-1}}\right)$$

where $\alpha = 0.97650403$

$\beta = 0.93112947$

$\gamma = 0.99700449$

The discrete-time step response of the plant with the compensator is shown in Figure 4.1 and it behaves exactly like the continuous-time compensated system. This is the discrete-time compensator (4.2) that would be ideally implemented. It is clear that the bilinear transform preserved all of the desired characteristics of the compensator.

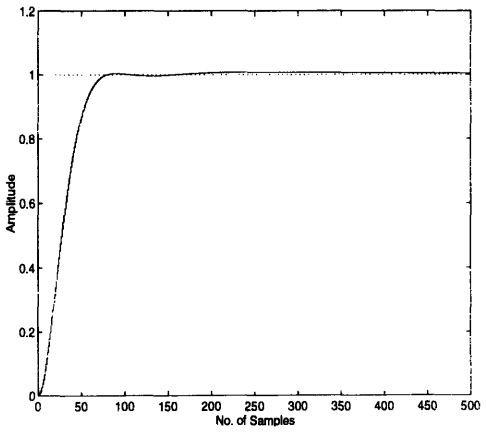


Figure 4.1: Step response of DT compensator derived by the bilinear transform.

$T_{sample} = 10^{-2}$ seconds

4.2 Servo Motion Controller Card

Unfortunately, the ability to create this compensator exactly does not exist since an off-the-shelf servo motor card with limited flexibility was previously chosen for this prototype application. A servo controller card was chosen for this project because of its capability to control up to four axes from a personal computer without troubling the user with the details of implementing the interface with both the microprocessor and the electromechanical system. Although the card limits the designer's flexibility for the sake of ease of use, the card provides functions with which not only to program coordinated motions of the four axes, but also to compensate the servos. A suitable compensator must then be designed using the tools provided by the card.

The controller card provides one particularly useful function: a "proportional, accumulator, difference" digital filter, the continuous-time equivalent of a PID (Proportional, Integral, Derivative) filter [17]. This function is a digital implementation of a lead-lag compensator. The difference equation that models this filter is

$$y[n] = \mathbf{K_p} e[n] + \mathbf{K_i} \frac{\sum_{m=0}^{\mathbf{lim}} e[n-m]}{256} + \mathbf{K_d}(e[n] - e[n-1]) \ , \tag{4.3}$$

where y[n] is the output, e[n] is the input (the difference between the desired and actual position), **lim** is the summation limit, and $\mathbf{K_p}$, $\mathbf{K_i}$, and $\mathbf{K_d}$ are proportional, "accumulator", and "difference" gains. The first term is the proportional gain term; setting all other variables to zero makes a compensator a simple proportional gain controller. By appropriately choosing $\mathbf{K_p}$, $\mathbf{K_i}$, $\mathbf{K_d}$, and **lim**, the dynamics of the compensator can be specified to suit the application. However, this is far from the form of the lead-lag compensator described previously. Considering the frequency domain equivalent representation of this equation makes this evident. The z-transform for (4.3) is

$$T[z] = \frac{Y[z]}{E[z]} = \mathbf{K_p} + \mathbf{K_i} \sum_{m=0}^{\mathbf{lim}} z^{-m} + \mathbf{K_d}(1 - z^{-1}) \ . \tag{4.4}$$

The middle term of (4.4) is the transform for the "integral" term, which would be more appropriately called a finite summation. That term can be re-written as

$$\mathbf{K_i} \sum_{m=0}^{\mathbf{lim}} z^{-m} = \frac{\mathbf{K_i}(1 - z^{-\mathbf{lim}})}{1 - z^{-1}} \tag{4.5}$$

which means that the transfer function (4.4) can be re-written as

$$T[z] = \mathbf{K_p} + \frac{\mathbf{K_i}(1 - z^{-\mathbf{lim}})}{1 - z^{-1}} + \mathbf{K_d}(1 - z^{-1}) \ . \tag{4.6}$$

It may seem like (4.5) is the DT version of the lag compensator desired because it has a pole at one. However, (4.5) has a zero at one since all roots of $1 - z^{-\mathbf{lim}}$ lie on the unit circle with the first root at one. This effectively cancels the pole at one, the pole that is desired. Qualitatively, this compensator does not provide infinite DC gain because it is only performing finite summation. For a constant DC (error) input, a true integrator will

produce a ramp to infinity, as shown in Figure 4.2. For the finite summation of (4.5), a

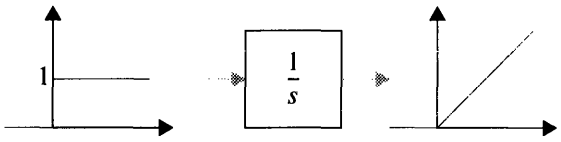


Figure 4.2: Response of an integrator to a step

constant input will ramp and reach a maximum. If the step input were a position error signal, then with a true integrator, the control signal output would grow to infinity in an attempt to reduce the error. In the controller card, the output would reach a limit. In the limit as **lim** goes to infinity, the $z^{-\mathbf{lim}}$ vanishes and (4.5) does becomes a true integrator. However, the **lim** is only allowed to take a finite maximum value [17]. Another practical difficulty that arises is that the card keeps numbers up to 16-bit precision, i.e. the largest value it can count up to is 65536. For values that attempt to go beyond that, it holds the maximum value. Thus, the integral controller differs from the ideal DT lag compensator in two ways: it does not perform infinite summation, and it is susceptible to variable saturation. The net effect is that it will not provide the pole at zero required.

Second, the compensator provides no provisions for pole placement. All poles are fixed at the origin according to the model. Furthermore, the only zero placement available is that of the derivative term. All other zeroes from the integral term are restricted to the unit circle. As described in chapter 3.0.3 for a lead-lag compensator, the pole and zero placement essentially defines the performance of the compensator. Restricting this ability limits the capabilities of this PID filter.

The solution to this control problem does not solely lie in selecting the appropriate values of $\mathbf{K_p}$, $\mathbf{K_i}$, $\mathbf{K_d}$, and **lim**. The user has been handicapped in tailoring the appropriate compensator for the bender module. First, the servo controller card must be taken as far as it can go. Since the integral compensator will not provide the pole at zero required, there is no reason to use it. The best must be made out of the PD compensator that is available. At that point, more satisfactory results can be found by considering the shape that the input takes.

4.3 PD compensator

The same method for the development of the continuous-time lead compensator can be applied to implement a discrete-time compensator, given the constraints of the card. That is, based on the open loop frequency response, determine what aspects of the frequency response needs to be adjusted, and choose the best possible **Kp** and $\mathbf{K_d}$ that comes closest to satisfying these specifications. The main difference is that instead of using the transfer function described in (3.6), the function in (4.4) was used.

As described in the continuous-time derivation, the three aspects that need to be improved were the low crossover frequency, the low phase margin, and the steady state error. Improving them would bring the step response of the system toward what is desired: a minimum rise time of 0.5 seconds, and no overshoot, or oscillations.

Since the "integral" compensator will not provide the infinite DC gain required, there is no reason to use it. Given that $\mathbf{K_i}$ and **lim** are zero, the difference equation of the PD filter is

$$y[n] = \mathbf{K_p} e[n] + \mathbf{K_d}(e[n] - e[n-1]) . \tag{4.7}$$

The z-transform for this difference equation is consequently

$$Y[z] = \mathbf{K_p} E[z] + \mathbf{K_d}(E[z] - z^{-1}E[z])$$

$$H[z] = \frac{Y[z]}{E[z]} = \mathbf{K_p} + \mathbf{K_d}(1 - z^{-1}) . \tag{4.8}$$

The pole of this compensator is fixed at the origin. The zero of this compensator is located at $\frac{\mathbf{K_d}}{\mathbf{K_p} + \mathbf{K_d}}$ and can take any value between 0 and 1 depending on the choice of $\mathbf{K_p}$ and $\mathbf{K_d}$. $\mathbf{K_p}$ and $\mathbf{K_d}$ can be any integer value between 0 and 25000.

A first attempt at specifying the PD controller would be to try to implement the ideal PD compensator for this system as described in chapter 4.1. To do this, $\mathbf{K_p}$ and $\mathbf{K_d}$ must be chosen such that the zero is located at the same place as the zero in the ideal PD compensator for sampling period of $T_{sample} = 10^{-4}$ seconds: $\Omega_z = 0.999762275$ radians[1]. However, the pole is fixed at 0 radians instead of $\Omega_p = 0.999286987$ radians, where it should be

for the PD compensator with a sampling period of $T_{sample} = 10^{-4}$ seconds. This is a clear indicator that the exact PD controller cannot be implemented. Although a pole at the origin is equivalent to a pole at negative infinity in continuous-time and by itself has no effect on the frequency response of a system, the pole of a pole-zero pair of the PD compensator together has an important impact on the dynamics of the system. As previously mentioned, the zero serves to increase the phase of the frequency response, but it also increases the magnitude with increasing ω. This effect is generally undesirable since it makes the system sensitive to high frequency inputs and even potentially unstable. The pole serves to localize the compensation to a small range of frequencies by "undoing" what the zero did. Therefore, without the pole to complement the zero, caution must be used in the selection of the proportional gain and the placement of the zero.

If the zero was placed at $\Omega_z = 0.999762275$ radians, the primary concern would then be what gain to use. Upon examination of the open-loop frequency response of the plant, it is not a monotonically decreasing function due to the peak attributable to the complex-conjugate pole pair at $z = 9.9418\times10^{-1} \pm 2.5854\times10^{-2}j$ (or equivalently at $s = 55 + 260j$). As such, the proportional gain cannot be so large as to create multiple crossover points. The danger of creating multiple crossover points is even greater when the rate of attenuation of the frequency response is decreased by the addition of a zero at 0.999762275 radians.

---

1. For $T_{sample} = 10^{-2}$ seconds, the pole and zero pair for the lead compensator found by applying the bilinear transform is $\Omega_p = 0.93112947$ radians and $\Omega_z = 0.99700449$ radians. However, for $T_{sample} = 10^{-4}$ seconds, the pole and zero pair is $\Omega_p = 0.999286987$ radians and $\Omega_z = 0.999762275$ radians. They are equivalent given the different sampling periods - their derivations are by the same process. Although using $T_{sample} = 10^{-4}$ seconds would have been more consistent with the entire chapter, it was not possible to simulate the compensator found by the bilinear transform (chapter 4.1) on Matlab with that sampling period. Many values were close to being singular.

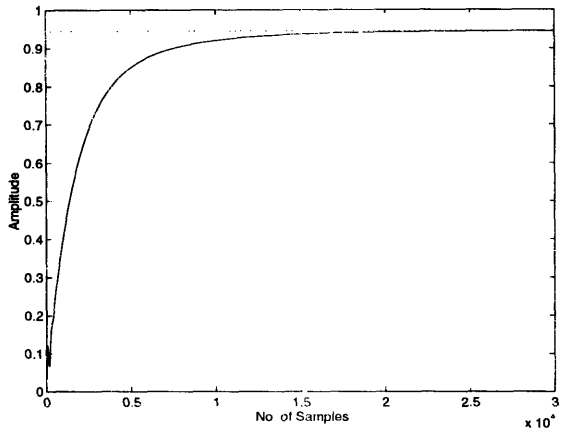


Figure 4.3: Step response of PD compensated system using K_p=5.43 and K_d=25000.

$T_{sample} = 10^{-4}$ seconds

$\mathbf{K}_p$ = 5.43 and $\mathbf{K}_d$ = 25000 places the zero at $W_z = 0.999762275$. The step response is shown in Figure 4.3. The rise time is very good, but without the infinite DC gain, the steady state error of the system is non-zero.

An alternative to using a "integral" controller to provide open-loop infinite DC gain is to use a very large proportional gain. Again, consider the error function discussed in chapter 2.3 where *P(s)* denotes the transfer function of the plant and $\mathbf{K_p}$ is a proportional gain

$$\frac{E(s)}{R(s)} = \frac{1}{1 + KP(s)}. \tag{4.9}$$

As $\mathbf{K_p}$ approaches infinity, the steady state error approaches 0. Thus, for practical purposes, a high proportional gain will be sufficient to reduce the error to a value that is approximately zero. However, as also discussed, a high proportional gain will reduce the phase margin of the system since it changes the crossover frequency without changing the phase response of the bender unit. When the phase margin is zero, the system is on the verge of instability.

Since a large $\mathbf{K_d}$ keeps the zero close to the region that would benefit from the positive phase contribution, choose $\mathbf{K_d}$ to have the maximum value at 25,000. Choosing a large $\mathbf{K_p}$ around 2000 will result in zero steady state error in the step response.

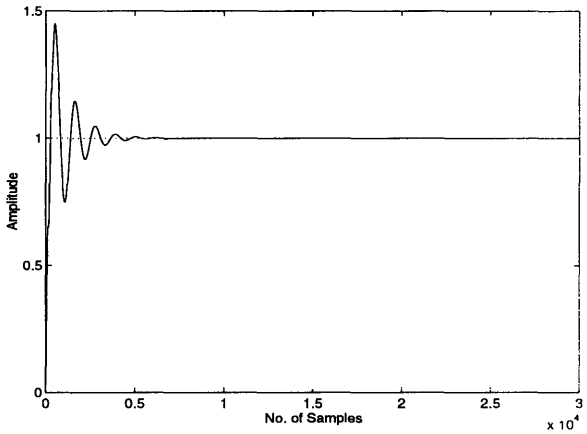


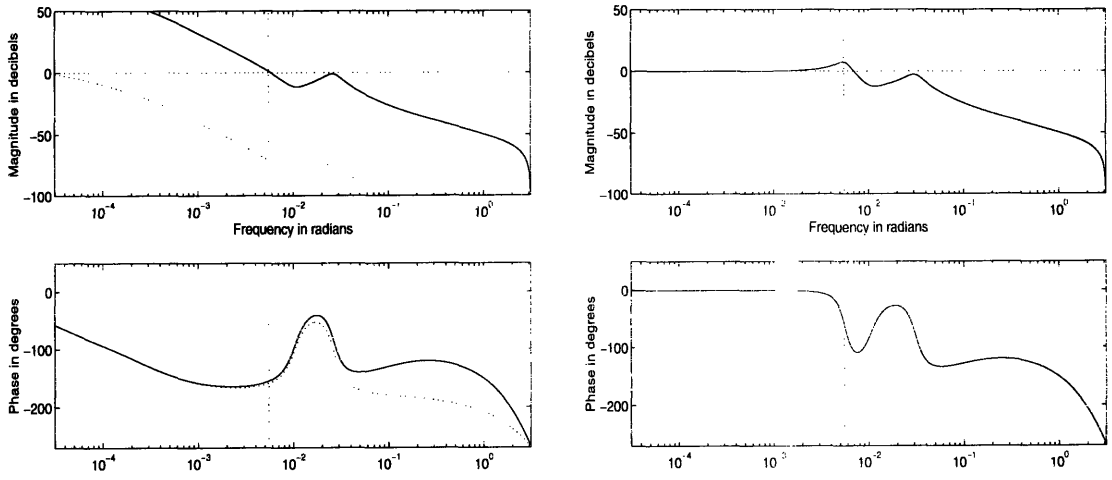
Figure 4.4: Step response of PD compensated system using K_p= 2000 and K_d=25000.

$T_{sample} = 10^{-4}$ seconds

The step response as shown in Figure 4.4 has a short rise time and zero steady state error, but unfortunately it overshoots and oscillates badly. On the open-loop frequency response, this should be evident based on the fact that the system has very low phase margin - roughly 23° - as shown in Figure 4.5. On the closed-frequency response, there exists a peak at $\omega = 5.5175\times10^{-3}$ radians which is due to the low damping in the system, also shown in Figure 4.5. Signals with a spectral content in the range of the peak are amplified with a gain greater than unity. Thus, for a step input - a signal which can be thought of as having a Fourier expansion of an infinite sum of sinusoidal waveforms [13] - the system

Figure 4.5: Open-loop and closed loop frequency responses PD compensated system K_p = 2000 and K_d = 25000.

The phase margin on the open-loop response (left) is 23°. Peaking occurs on the closed-loop response (right) is at 5.5175×10^{-3} radians.



will amplify components with those frequencies. The step input excites the system at its resonant frequency. The result is the ringing.

This insight can be used to achieve the desired response. First of all, the real system will never be driven by a step input. A step in angular position corresponds to an impulse in angular velocity and a doublet in acceleration. That is, a step command requires an infinite positive velocity at time zero, and instantaneous infinite positive and negative acceleration at time zero. Since acceleration is proportional to the torque by Newton's second law of motion, it implies that a step in position requires an infinite amount of torque in the positive and negative direction at time zero. Since the amplifier that drives the motor has a current limit, the torque that the motor can supply is limited by the maximum output current from the amplifier.

The real system will require a finite torque over a finite time interval. The system will accelerate, then decelerate to the final position. Given the flexibility in choosing the input that will bring the system to its final position, choose one that will not excite the resonant frequency at $\omega = 5.5175\times10^{-3}$ radians. A waveform that is purely sinusoidal has one frequency. If that frequency was chosen to be far below the peaking frequency, then the output of the system driven by an input of that frequency will be in phase and equal in magnitude with that input. In Figure 4.6, the PD compensated system is driven with an 0.5 Hz sinusoidal input. The signal is in phase and equal to the input. Some transient ripples exist in the beginning of the output due to the instantaneous change in velocity at time

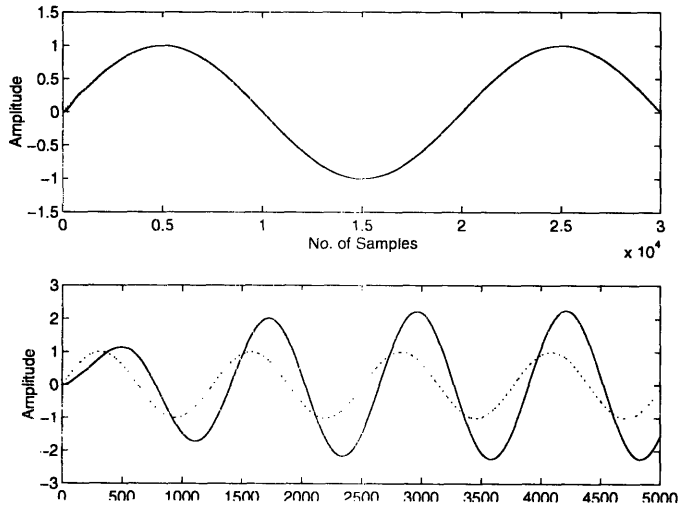


Figure 4.6: PD compensated system, K_p= 2000 and K_d=25000, driven by sinusoidal inputs

Top: 0.5 Hz (3.14×10^{-4} radians) sinusoidal input. System output (solid line) is in phase and equal in magnitude.

Bottom: 7.95 Hz (5×10^{-3} radians) sinusoidal input. System output (solid line) depicts slight phase shift and amplification relative to the input (dotted line).

$T_{sample} = 10^{-4}$ seconds

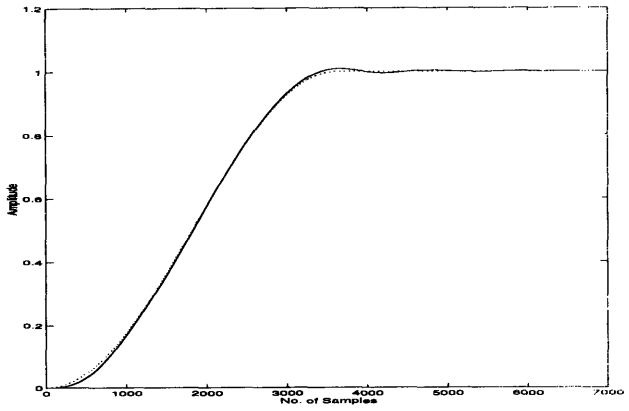


Figure 4.7: Response of PD compensated system to half period of a sinusoidal input.

The response (solid line) tracks the input (dotted line) very well to half period of sinusoidal input with $\omega = 2.75\pi\times10^{-4}$ radians and 90° phase shift.

$T_{sample} = 10^{-4}$ seconds

zero. With a 7.95 Hz sinusoidal input, the output is no longer in phase and the amplitude no longer unity, as expected from the closed-looped frequency response.

An input can be created that utilizes this concept. If the step was actually half of a period of a sinusoid with a 90° phase shift and a DC offset of one-half, the input will look like Figure 4.7. If the sinusoid had a frequency of $2.75\pi\times10^{-4}$ radians, or 1.375 Hertz, then this will be sufficient to not excite any resonant frequencies or overshoot, as shown in Figure 4.7. Furthermore, the rise time specification is met, $n_{rise} = 3636$ samples or $t_{rise} = 0.36$ seconds. The overshoot and steady state error is zero.

4.4 Experimental Verification

To verify that the final controller implementation complemented with input shaping works, a set of experiments were conducted using the servo controller card. The PID filter values were set as stated in the previous section ($K_p = 2000$, $K_d = 25000$). The input was shaped not exactly as a sinusoid with a DC offset, but as a waveform with a trapezoidal velocity profile. That is, the input signal was created by specifying an acceleration, a maximum velocity, and a desired position. From zero position, the motor was commanded to accelerate at *a*, continue at velocity *v*, and then decelerate at *a* to reach the final position *x*.

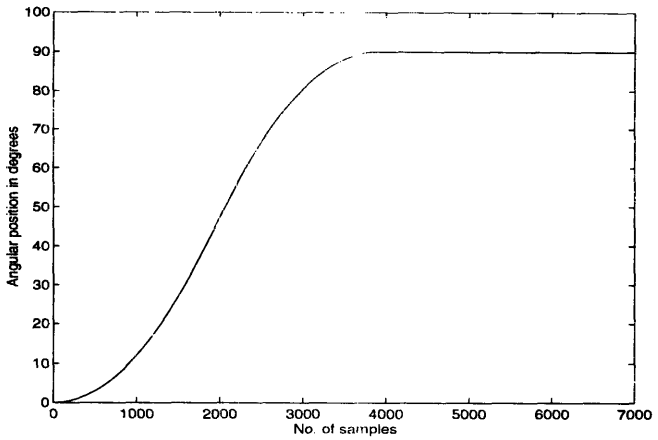


Figure 4.8: Waveform generated by using a trapezoidal velocity profile.

$x = 90^{o}$

$v_{max} = 7.34$ radians/second

$a = 41.95$ radians/second^{2x}

$T_{sample} = 10^{-4}$ seconds

The resulting waveform is approximately sinusoidal, as shown in Figure 4.8. A Matlab function was written to generate such profiles, using an acceleration a, a maximum velocity v, and final position x. It can be found in the [11].

A 2.75π radian/second sinusoidal waveform was approximated by a waveform with a trapezoidal velocity profile with an acceleration of 41.95 radians/second^2 and a maximum velocity of 7.34 radians/second. Using these parameters and a $\frac{\pi}{2}$ radians (90^{o}) desired position, the input profile is shown in Figure 4.8. The simulated and actual response of the system is shown in Figure 4.9. It verifies that compensation with input shaping achieves the desired requirements of fast response (with a rise time to 90^{o} of 0.35 seconds), zero steady state error, and no overshoot.

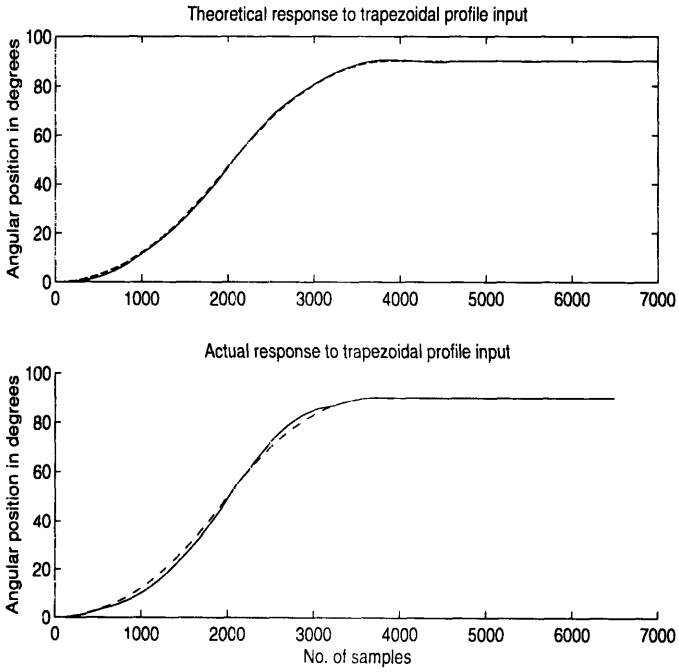


Figure 4.9: Theoretical and actual response of PD compensated system to a trapezoidal velocity profile input.

$T_{sample} = 10^{-4}$ seconds

5 Bending Algorithm

The ultimate goal of the bender module is to be able to bend automatically. That is, the bender module must be able to bend parts to any specified angle between 0 and 90°. The bending process must be repeatable and accurate to within ±0.5°. To achieve this ultimate goal, the bender must be able to compensate for material characteristic variability of the parts. It must be able to sense differences in the actual angle versus the desired bend angle. Furthermore, it must be capable of making minute changes or corrections of the bend angle of the parts.

5.1 Description of Physical System

In order to deal with these challenges, it is necessary to understand the context of the problem, i.e. the physical realization of the bender [2] and the resources available. As briefly described in Chapter 1, the bender module is essentially a DC servo coupled to an "arm." This "arm" is an aluminum U-frame with a pair of grippers positioned at the center of it. More specifically, the foremost edges of the grippers are aligned onto a plane that contains the axis of rotation. The gripper surfaces are also always equidistant from the axis of rotation and parallel to each other, whether the grippers are opened or closed. These grippers are spring-loaded in a normally open position. A pneumatic piston, mounted behind the grippers on the U-frame, exerts the force necessary to close them. Since the piston acts in a direction orthogonal to the direction of the desired motion, the force is translated through wedges mounted as shown in Figure 2.1. When the piston is extended, the grippers are closed; when it is retracted, they are opened.

The means to create and execute a bending algorithm is a personal computer. This computer controls the servo motor via a servo control card. The PD controller as well as the input shaping scheme has been implemented as described in Chapter 3. The computer also opens and closes the piston through a single binary control signal. All that is required is the software program to manipulate the motor and the pneumatic piston to perform bends.

5.2 Elastic Springback and Material Characteristic Variability

The biggest challenge in creating a successful bending algorithm is to understand how to deal with the variable material characteristics of different parts. Specifically, the characteristic that is of great importance is the elastic springback of the part after it has been bent. An understanding of what gives rise to springback is useful for later derivation of a bending algorithm.

The amount of springback depends on the elastic-plastic characteristic of the metal from which the part is made [4]. Think of the part as a slender metal bar with uniform material characteristics. When a pure bending moment or torque is applied to the part, it will at first elastically deform. In this region, when the moment is removed, the metal will revert to its original position. The curvature ρ^{-1}, defined as the inverse of the radius of the bend, ρ, and the moment applied, M, will vary linearly. As depicted in a typical moment-curvature curve in Figure 5.1, it will traverse segment OA. At some moment M_y, the mate-

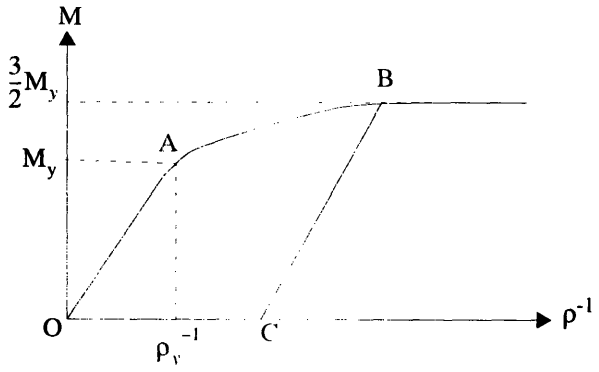


Figure 5.1: Moment-curvature relationship

rial yields; that is, instead of elastically deforming, it will begin to plastically - and permanently - deform. The moment-curvature relationship is no longer linear, and as derived in [4], the curvature increases to infinity as the moment increases to an asymptotic value of $\frac{3}{2}M_y$ - the fully plastic moment. The moment-curvature relationship depends on the material composition of the part as well as the dimension. Thus, the variability can be great between different types of parts.

When a bending moment that plastically deforms the part is removed, the metal undergoes elastic recovery. As the moment decreases to zero, the resultant curvature decreases linearly along the line BC parallel to OA. This change in the curvature is called elastic springback [4].

Elastic springback makes bending to small tolerances difficult, especially when the amount of springback can vary across parts of different composition. Bending to a desired angle can be done in one attempt only if the overbend angle required to compensate for springback is known. This implies that the moment-curvature relationship for the particular part being bent must be known well enough so that the required overbend angle can be determined. Without exact knowledge of the relationship or the overbend angle, multiple attempts must be made to bend a part to a desired angle. Essentially, the bending algorithms developed in the following sections consist of iterative methods of finding the right overbend angle with which to achieve the desired angle. This was accomplished using no knowledge of the moment-curvature relationship (chapter 5.4) and later with partial knowledge (chapter 5.5).

5.3 Angle Measurement by Electrical Continuity

Given the current physical realization of the module, several methods of measuring the actual unloaded angle of the part - i.e. with no moment applied - after it is bent were considered [2]. The method that was determined to have the best chance of working well utilized the conductivity and springback of the parts and the encoder available on the bender module.

Since the angular position of the bender arm is always known (from monitoring the encoder mounted on the motor) and in very close proximity to the bent parts, it is worthwhile to try to measure unloaded bend angle relative to the angular position of the bender arm. With the grippers opened slightly, the gripper and the segment of the parts inserted in

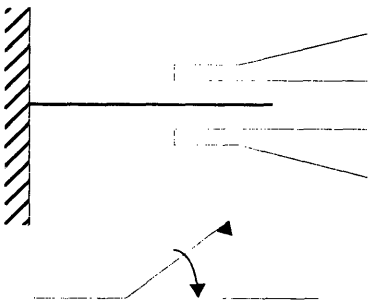


Figure 5.2: Grippers and part considered as a switch

the grippers can be thought of as a switch, as shown in Figure 5.2. After a bend is made and the grippers opened, the natural springback that the part is assumed to have will keep the switch "closed". Of course, this implies that the gap be small enough so that the springback of the pin is enough to maintain contact with one of the grippers. Conceptually, if the arm was rotated back in the direction of the elastic moment of the part, the angle of the unloaded part is approximately the angle at which time the switch opens. Thus, the electrical continuity through the part and the gripper in conjunction with the readings from the encoder can be utilized to measure the unloaded bend angle of the parts. By "pulling" the part up to 5 volts and the grippers to ground as shown in Figure 5.3, the resulting

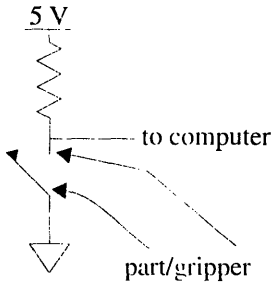


Figure 5.3: Pullup circuit for continuity measurements.

binary electrical signal can be monitored by a computer to determine continuity.

The angle measured is only a very close approximation since the gap between the grippers introduces an error in the measurement of the unloaded angle. Since the angle measured is the angle of the grippers when continuity is lost with the part, the error, ε, is

$$\varepsilon = \Theta_{actual} - \Theta_{measured} = \text{asin}\left(\frac{0.5g}{l}\right) \tag{5.1}$$

where Θ_{actual} and $\Theta_{measured}$ is the actual and measured bend angles, g is the gap between the grippers, and l is the length of the segment within the grippers as shown in Figure 5.4.

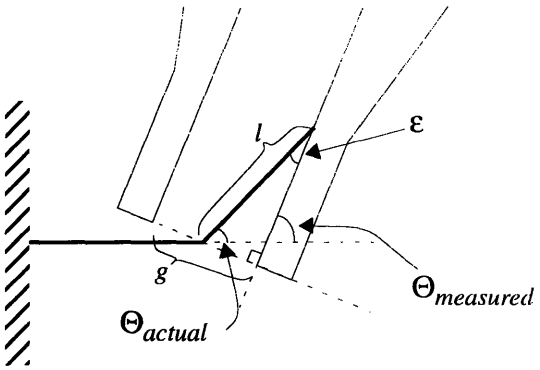


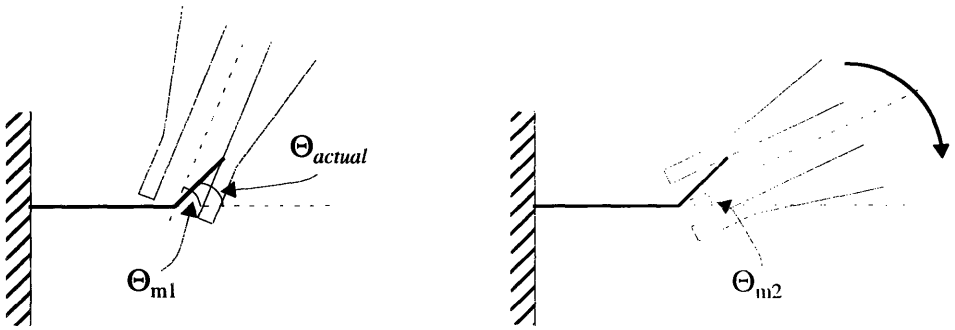
Figure 5.4: Measurement error in using continuity with the grippers opened.
Θ_{actual} and $\Theta_{measured}$ is the actual and measured bend angles.
g is the gap between the grippers
l is the length of the segment within the grippers

Although g is measurable and fixed, l is difficult to measure and varies depending on the length of the part to be bent that is inserted into the grippers. This length varies for the same part in the case where multiple bends are performed on it.

A solution to this problem is to make two measurements, one when continuity is lost and one when continuity is regained, and use their average. As the opened grippers rotate in the direction of the springback as shown in Figure 5.5, the measurement when continuity is first lost, Θ_{m1}, has error ε whereas the measurement when continuity is regained,

Figure 5.5: Angle measurement concept
Two measurements, Θ_{m1} and Θ_{m2}, are made as the grippers rotate. The average is the actual angle, i.e.

$$\Theta_{actual} = \frac{\Theta_{m1} + \Theta_{m2}}{2}$$



Θ_{m2}, has error $-\varepsilon$. The average of the two measurements theoretically removes any error in the final measurement used. The actual angle, thus, is

$$\Theta_{actual} = \frac{\Theta_{m1} + \Theta_{m2}}{2} \tag{5.2}$$

This method was implemented and a set of tests were conducted to prove its validity. The bender was commanded to bend a part to a 90° angle, and then measure the resulting springback angle. This measurement was compared to a measurement of the same part using a Vernier protractor. Table 5.1 lists the measurements and the errors. The set of tests

Desired angle	Springback measured angle by Bender	Springback measured angle by Protractor	Error
90	65.526	66.00	0.974
90	64.560	65.50	0.940
90	64.560	65.50	0.940
90	64.845	65.92	1.075
90	65.329	66.33	1.001
90	64.384	65.25	0.866
90	65.043	66.00	0.957
90	65.131	66.09	0.959
90	64.472	65.42	0.948
90	65.153	66.00	0.847

Table 5.1: Comparison of springback angles measured by Bender and by Vernier protractor. All ten tests were made to the same angle. The average error is 0.95° and the standard deviation of the error is 0.064°.

clearly indicates an average error of 0.95° which was the result of misalignment of the plane of reference for the part and the bender arm. The standard deviation of the errors was 0.064° indicating a high degree of consistency and precision in this method.

5.4 Algorithm Using Proportion of Error

Given the means of measuring the unloaded bend angle of the parts, the simplest algorithm is to iteratively check and bend the part until it is has the desired bend angle. After a bend is performed, measure the actual angle. If the actual angle does not fall within a specified tolerance band, increase or decrease the previously commanded bend angle by some proportion of the error. This method attempts to compensate for the unknown elasticity of the metal by overbending the part until the unloaded angle equals the desired bend angle. Given the i^{th} attempt, the commanded bend angle, Θ_c, at $i+1^{th}$ attempt is defined by

$$\Theta_c[i+1] = \Theta_c[i] + b\varepsilon_\Theta[i] \tag{5.3}$$

where the original commanded position is

$$\Theta_c[0] = \Theta_{desired} \tag{5.4}$$

The proportionality constant, b, is largely defined by the requirements for speed of bending a part to a certain angle. It must be large enough so that a small reasonable number of iterations are required. If the constant was very small, the step changes in $\Theta_c[i+1]$ would be so small that convergence to the actual position to the desired one would be slow, i.e. require several iterations. On the other hand, b must not be too large so as to cause the actual bend angle from exceeding the desired angle. "Overshooting" the desired bend angle requires bending the part in the opposite direction. Repeated bends back and forth will tend to strain-harden the metal which is undesirable [4]. Thus, the constant b should be chosen carefully such that the iterative bending process increases the actual bend angle towards the desired bend angle quickly and monotonically.

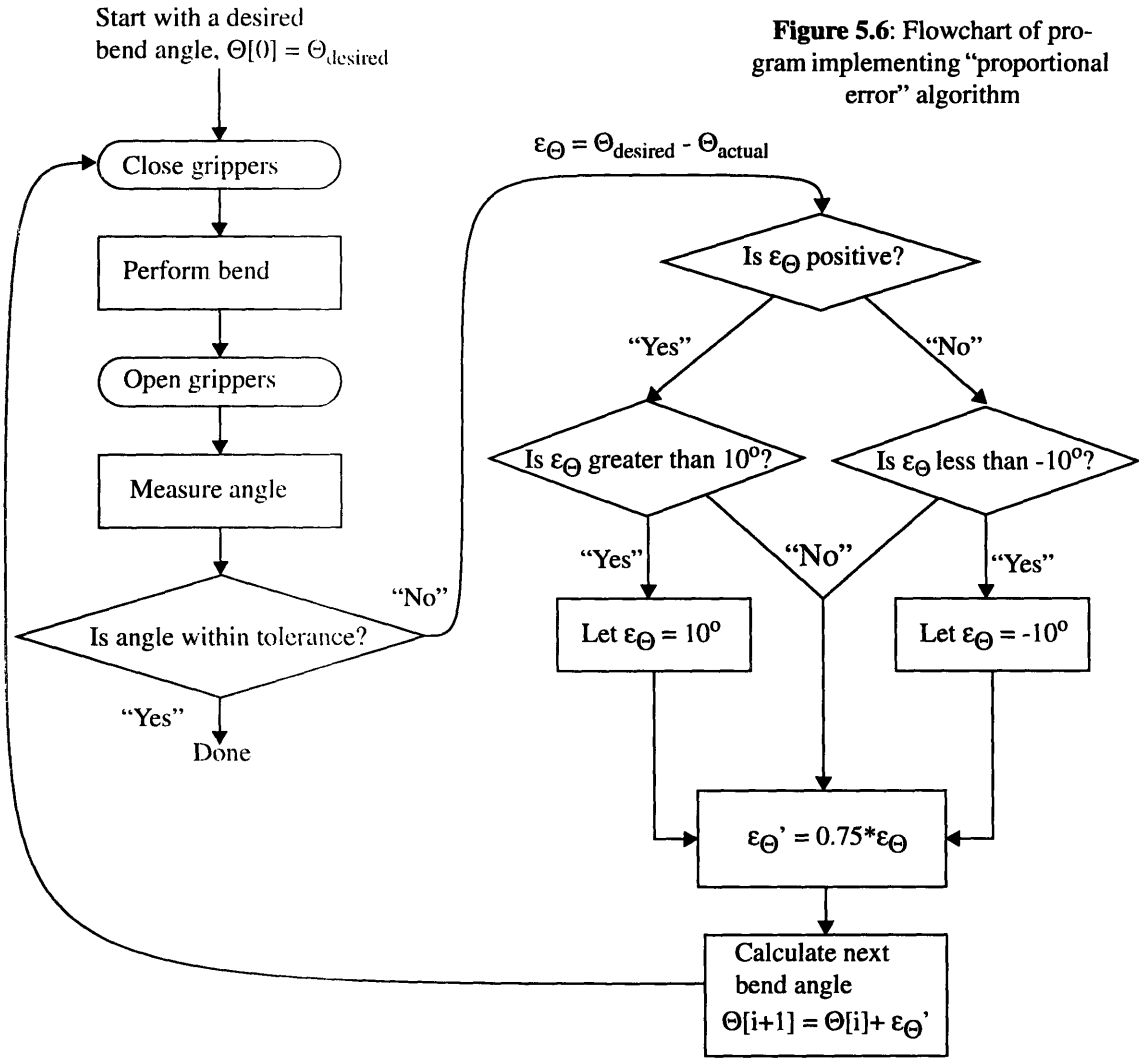


Figure 5.6: Flowchart of program implementing "proportional error" algorithm

5.4.1 Implementation

A program was written in C to implement this algorithm. Figure 5.6 depicts a flow chart of the relevant procedures in the implementation. The first bend is made at the desired bend angle. If there is no springback and the bend angle is within the specified tolerance, then the program exits the loop. Otherwise, the program iteratively makes improved guesses on the right bend angle to achieve the desired unloaded angle. In general with each iteration, the next bend angle is equal to the previous bend angle plus a proportion of the error. But if the magnitude of the error exceeds the 10°, then a maximum value of 10° is used, and a proportion of that is added to the previous bend angle. This maximum error was arbitrarily chosen and implemented as a precaution against inadvert-

ently overshooting the desired bend angle. It also allows the use of a larger b for faster convergence to the desired angle. This maximum error limit, the proportionality constant as well as the tolerance band are variables that can be arbitrarily chosen. The program can be found in the "Flexible Assembly Project Software Technical Report" [11].

5.4.2 Experiments

Bend experiments, documented in [11], were performed on 0.5" by 1.5" sheets of stock material from which parts are typically made. They were more economical to use and served the same function as the parts. A tolerance of $\pm 0.5^\circ$ was specified, and experimentally, it was found that $b = 0.75$ allowed the bender module to converge to the desired angle in 4 or 5 iterations for this sample material. The maximum error limit of 10° was required so that the program would not overshoot the desired bend angle. Figure 5.7

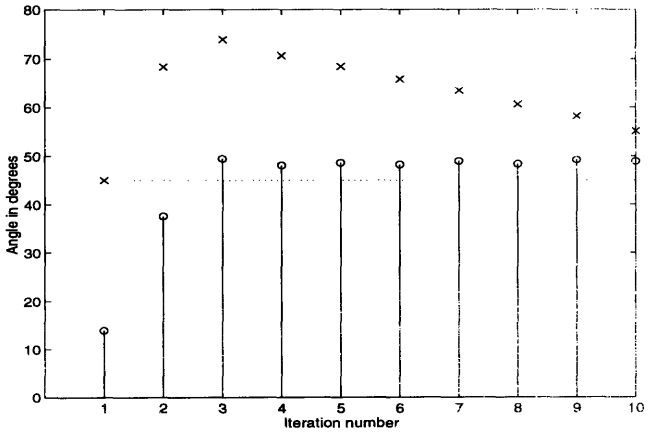


Figure 5.7: Results of 45° bend experiment where maximum error bound was not used

'x's denote the overbend angle used

'Stems' denote the bend actual bend desired, while the dotted line marks the desired bend angle.

Note how the bender attempts to compensate by decreasing the overbend angle. Refer to Table 5.2 for numerical values.

shows the results of one experiment where the maximum error bound was not used in the algorithm. The actual and overbend angles as well as the error ε were recorded for a 45° desired bend and listed in Table 5.2. Without the error bound, the algorithm causes the bender to overshoot the desired bend angle. Within ten iterations, it is not able to recover and bend the part to 45°.

step	overbend angle	actual angle	error, ε	time
1	45.00	13.93	31.07	2.42
2	68.36	37.58	7.42	5.00
3	73.97	49.43	-4.43	7.30
4	70.68	48.06	-3.06	9.45
5	68.46	48.60	-3.60	11.53
6	65.84	48.22	-3.22	13.57
7	63.50	48.94	-3.94	15.60
8	60.63	48.33	-3.33	17.47
9	58.20	49.19	-4.19	19.22
10	55.13	48.87	-3.87	20.87

Table 5.2: 45° bend using "proportional error" algorithm, without maximum error bound. Measurements at each step are recorded. The overbend and actual angles and the error are measured in degrees. Time durations were measured in seconds.

Figure 5.8 shows the results of the same 45° bend experiment, but this time, the maximum error bound of 10° has been incorporated in the program. Table 5.3 lists the numerical values from the experiment. The plot of the bend angle clearly shows where the maximum error limit has effectively limited the increase. For the second and third bends, the overbend angle was limited to the previous angle plus 7.5°. The results of more experiments that were conducted in the same manner at different angles using this algorithm can be found in [11].

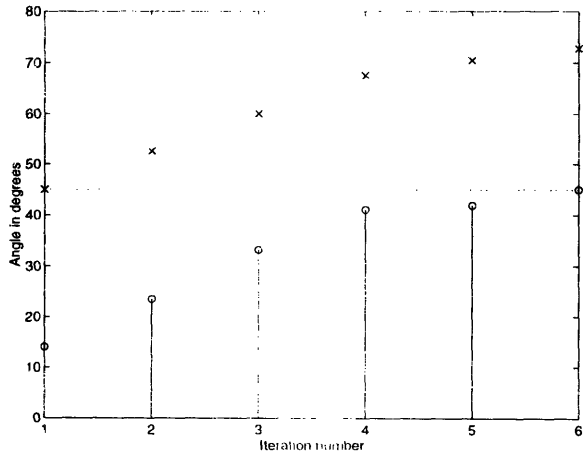


Figure 5.8: Results of 45° bend experiment where maximum error bound was used

Note that the change in overbend angle is linear for the first four iterations. The maximum error limit prevents the change in overbend ('x's) angle per iteration from being greater that 7.5°.

The result is that it does not overshoot. Refer to Table 5.3 for the numerical values.

step	overbend angle	actual angle	error	time
1	45.00	14.02	30.98	2.12
2	52.50	23.54	21.46	4.39
3	60.00	33.19	11.81	7.25
4	67.50	41.11	3.89	9.56
5	70.45	41.94	3.06	11.92
6	73.82	45.05	-0.05	14.28

Table 5.3: 45° bend using "proportional error" algorithm, with maximum error bound
Measurements at each step are recorded. The overbend and actual angles and the error are measured in degrees. Time durations were measured in seconds.

Three sets of ten bends were also conducted. The purpose of these tests was to determine the accuracy and repeatability of the bender. For each set, the bender module was instructed to bend the parts to one of three angles: 30°, 60°, or 90°. The final bend angle that the bender module reached was recorded, and afterwards, the bend of each sample was measured by hand using a Vernier protractor. The "Flexible Assembly Project Software Technical Report" [11] contains the data gathered from the experiment. Table 5.4 contains a listing of relevant statistics for the three trials. There is a bias between the measurements made by the bender and those made by hand using a protractor. This is attributable to the misalignment of the plane of reference of the part and the bender arm. The larger standard deviation in the measurements was due to the limited resolution of the pro-

desired angle	Average angle measured by bender	standard deviation (bender)	Average angle measured by protractor	standard deviation (protractor)
30	29.69	0.0484	30.10	0.2104
60	59.60	0.1289	59.98	0.1784
90	89.81	0.0805	91.00	0.1621

Table 5.4: Relevant statistics from three sets of ten bends
The bender was instructed to perform three sets of 10 bends to one of 3 angles: 30°, 60°, and 90°. This table contains the average final angle for each set as measured by the bender and checked by using a Vernier protractor. The standard deviations are also listed.

tractor to 0.083° increments. In any case, the standard deviation of the measurements are well within the 0.5° tolerance band that was specified.

5.5 Algorithm Using Moment-curvature Relationship

The previous method attempts to compensate for springback without any knowledge of the characteristics of the metal. Of course, a better one can be found if elastic properties of the metal were taken into consideration. As described in 4.3, springback is due to the elastic recovery of a metal when the applied bending moment is removed. The moment-curvature curve of Figure 5.1 depicts the effect that springback has on the curvature of a beam. It also provides insight as to how to compensate for springback and arrive at a desired curvature.

Suppose that the parameter of interest and the one that can be measured is the curvature. Later the following derivation will be related to the actual parameter of interested: bend angle. Also assume that the metal is elastic but perfectly plastic, as shown in Figure 5.9. That is, in its plastic region, a constant applied moment will produce any curvature.

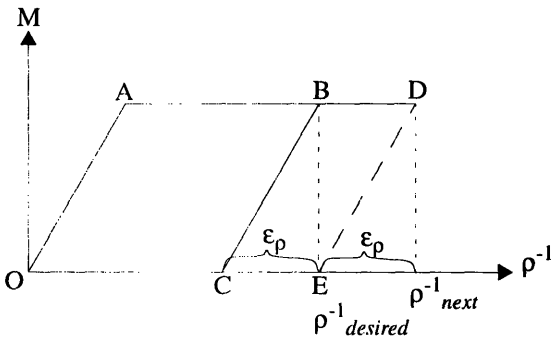


Figure 5.9: Moment-curvature relation for an elastic - perfectly plastic material

For this material, bending to $\rho^{-1}_{next} = \rho^{-1}_{desired} + \varepsilon_\rho$ will compensate for the springback of the material.

Since the part will springback after a bend to a curvature of $\rho^{-1}_{desired}$, the curvature, ρ^{-1}, after elastic recovery will be

$$\rho^{-1} = \rho^{-1}_{desired} - \varepsilon_\rho \tag{5.5}$$

where ε_ρ is the difference in curvature. Consequently, for a perfectly plastic metal, the next moment that is applied should result in a curvature of

$$\rho^{-1}_{next} = \rho^{-1}_{desired} + \varepsilon_\rho \tag{5.6}$$

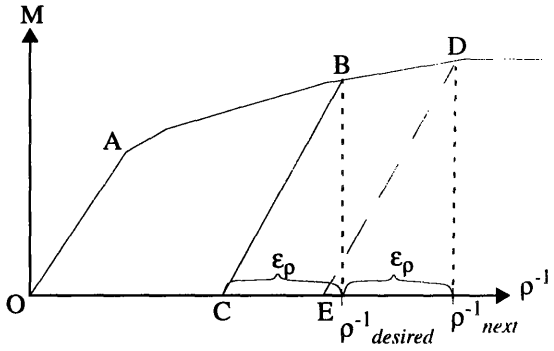


Figure 5.10: Moment-curvature relation for an elastic - **not** perfectly plastic metal

For this material, bending to $\rho^{-1}_{next} = \rho^{-1}_{desired} + \varepsilon_\rho$ will bring the curvature to a point E, just short of $\rho^{-1}_{desired}$.

so that when the moment is removed and the part elastically recovers, the part will have a curvature of $\rho^{-1}_{desired}$.

For metal that is not perfectly plastic, the moment-curvature relation will always have a non-negative slope. More specifically, the slope will be positive but asymptotically approach zero. In this case, specifying a moment that results in the curvature, ρ^{-1}_{next}, will result in a curvature that is just short of the desired one, as shown in Figure 5.10. Thus, the difference between loaded and unloaded curvature is a conservative value to use since it will never cause the bender to overshoot the desired curvature. Using the difference in curvatures as a correction factor seems like a viable proposal since it will produce a curvature that is at most equal to the desired one. As in the previous algorithm, repeating this process will result in a curvature that asymptotically approaches the desired. Thus given the i^{th} attempt, the next commanded curvature that the bender should attempt, $\rho^{-1}[i+1]$, is

$$\rho^{-1}[i+1] = \rho^{-1}[i] + \varepsilon_\rho[i] \tag{5.7}$$

where $\varepsilon_\rho[i]$ is the difference between the desired and actual curvatures after attempt i

$$\varepsilon_\rho[i] = \rho^{-1}_{desired} - \rho^{-1}_{actual}[i] \tag{5.8}$$

and the initial condition at the first attempt is

$$\rho^{-1}[0] = \rho^{-1}_{desired}. \tag{5.9}$$

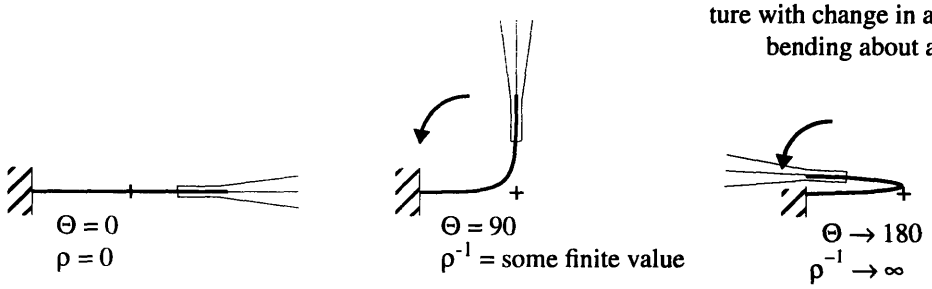


Figure 5.11: Change in curvature with change in angle during bending about an axis

However, the parameter of interest is the bend angle, which is non-linearly related to the curvature. This is evident by considering three cases where a part is bent about an axis of rotation by the gripper, as shown in Figure 5.11. If the part has a bend angle of $0°$, then the curvature is also 0, since this case corresponds to a bend around a surface of infinite radius. If the part has a bend angle of $90°$, then the curvature is some finite value and assumed to be non-zero. That is, the bend is not a sharp bend about a point. As the bend angle approaches $180°$, the curvature will approach infinity. These are the characteristics of a tangent function: zero at zero angle, finite as some angle, α, and infinite at twice that angle, 2α. Consequently, the non-linear relation between curvature and angle can be expressed as

$$\rho^{-1} = r^{-1} \cdot \tan\left(\frac{\Theta}{2}\right) \tag{5.10}$$

where r^{-1} is a scale factor. Its value is not important in the final result below, but included here for completeness.

With (5.10), (5.7) can be translated from an relationship between curvatures to one relating bend angles. Substitution of (5.10) for all curvatures in (5.7) gives

$$r^{-1} \cdot \tan\left(\frac{\Theta[i+1]}{2}\right) = r^{-1} \cdot \left(\tan\left(\frac{\Theta[i]}{2}\right) + \tan\left(\frac{\varepsilon_\Theta[i]}{2}\right)\right) \tag{5.11}$$

where $\varepsilon_\Theta[i]$ is the difference at the i^{th} attempt between the desired and actual bend angles

$$\varepsilon_\Theta[i] = \Theta_{desired} - \Theta_{actual}[i] \quad . \tag{5.12}$$

Cancelling the scale factor on both sides of (5.11) and taking the arctangent on both sides gives

$$\Theta[i+1] = 2 \cdot \text{atan}\left(\tan\left(\frac{\Theta[i]}{2} \right) + \tan\left(\frac{\varepsilon_\Theta[i]}{2} \right) \right) \tag{5.13}$$

where the first bend is made to the desired angle

$$\Theta[0] = \Theta_{desired} \tag{5.14}$$

The above equations, (5.13), (5.12), and (5.14), are equivalent to (5.7), (5.8), and (5.9) respectively, except that the non-linear translations between angle and curvature are included.

5.5.1 Implementation

A program was written in C to implement this algorithm. Figure 5.12 depicts a flow

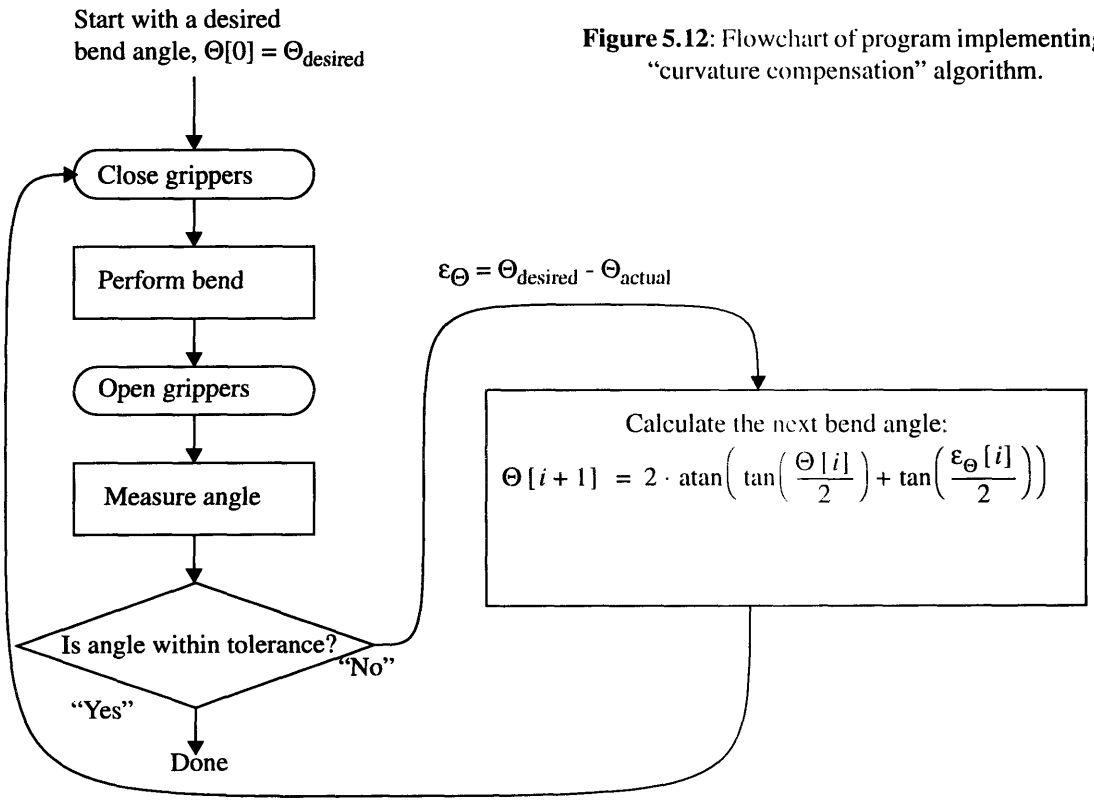


Figure 5.12: Flowchart of program implementing "curvature compensation" algorithm.

chart of the relevant procedures in the implementation. The first bend is made at the desired bend angle. If there is no springback and the bend angle is within the specified tolerance, then the program exits the loop. Otherwise, it calculates the next bend angle based on (5.11) and attempts another bend using the new bend angle. The program can be found in the "Flexible Assembly Project Software Technical Report" [11].

5.5.2 Experiments

Bend experiments, documented in [11], were repeated for this algorithm and compared with the results of the experiments using the "proportional error" algorithm. The same 0.5" by 1.85" samples of phosphor bronze were used, and a tolerance of ±0.5° was specified. Table 5.5 shows the results of one experiment where the actual and overbend angles as well as the difference, ε, at each iteration were recorded for a 45° desired bend angle. Table 5.5 lists the numerical values.

step	overbend angle	actual angle	error ε	time
1	45.00	13.91	31.09	4.31
2	69.84	40.71	4.29	4.67
3	73.26	44.35	0.65	6.97
4	73.89	44.93	0.07	9.23

Table 5.5: 45° bend using curvature compensation algorithm

Measurements at each step are recorded. The overbend and actual angles and the error are measured in degrees. Time durations were measured in seconds.

The most noteworthy aspect about the performance of this algorithm, as shown in Figure 5.13, was that no limit on the size of the error used for compensation was necessary. The virtue of this algorithm was that it can take large steps in the overbend angle without overshooting the desired bend angle. The theory of moment-curvature relationship of metals was used as the basis of this algorithm. However, the exact relationship for this particular alloy of the part was not used, nor was it necessary. No trial and error was necessary to fine tune the algorithm.

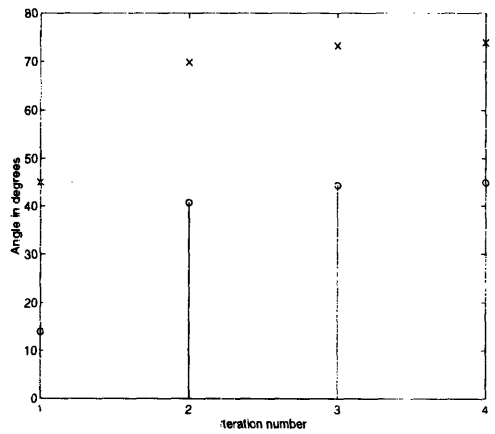


Figure 5.13: Result of 45° bend using 'curvature compensation" algorithm

6 Electromagnet for Clamping

The bender module is meant to work in conjunction with another module whose role is to transport a product to the bender and hold it securely while the bender operates. This transport module must be mobile so that it can transport the product in its possession to other modules. But since the product, the parts, and their tolerances are small, this transport module must be able to accurately and repeatably align itself with the bender module. It must securely dock with the bender module and remain stationary while the bender performs its task on the product. This chapter documents the design of a suitable clamping mechanism.

For accurate and repeatable docking, a kinematic coupling was chosen, as shown in Figure 6.1, a kinematic coupling is a connection where the mating surfaces are restricted to six points of contact [15]. Three equidistant hemispheres are built into one half of the

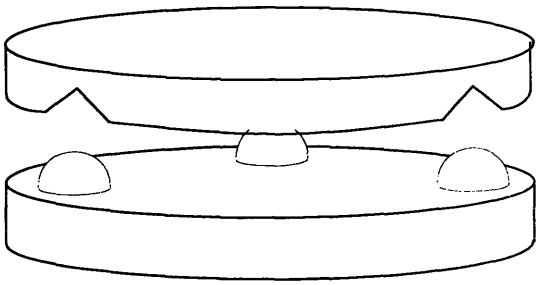


Figure 6.1: Kinematic Coupling

coupling, and three v-shaped grooves are machined out of the other half to mate with the hemispheres. The six degrees of freedom are constrained only at the point surfaces where the hemispheres contact the grooves. Repeatability of the relative position of the two halves can be as good as a few nanometers without requiring nanometer precision in the construction of the coupling.

To maintain the alignment of the transport module and the bender module through the use of the kinematic coupling, the two halves must be held together by a force that only acts in the direction that brings them together. Any lateral forces that exist can "unseat" the hemispheres from their v-groves and consequently, destroy the alignment of the two

halves. Thus, a clamp is required to apply a force in only one direction while not restricting motion in other directions.

Mechanical clamping techniques could be used for this application, but they add complexity to both the transport and bender modules. Strong specially shaped mating surfaces as well as additional moving parts may be required. A mechanical clamp would have to be designed so that it does not compromise the structural integrity of the modules or interfere with their operation. Finally, any mechanical clamping mechanism would also have to be carefully designed to prevent the application of lateral forces on the two halves of the coupling since this could destroy the alignment of the halves of the kinematic coupling.

An electromagnet is a simple alternative that can provide excellent clamping force in a small space. It requires no special mating surfaces, has no moving parts, and can be easily designed around components. An electromagnet is well suited for this clamping application because it provides force in only one direction - the direction that brings the two halves of the base together. It does not constrain motion in other directions or degrees of freedom. It, thus, satisfies a principal requirement for the clamp. Furthermore, since the base was originally designed to be machined out of iron - a good ferromagnetic material, the electromagnet can also be integrated into the design of the kinematic coupling. An electromagnet also has the useful feature that the force can be changed by adjusting the magnetizing current.

The requirements for this electromagnet that were specified for this application are:

1. It must provide a minimum of 100 pounds of compressive force on the two components of the kinematic coupling base.
2. It must not interfere with the operation of the kinematic coupling, i.e. exert no other forces on the halves of the coupling.
3. It must be capable of operating continuously.
4. It must be efficient so as not to require any special provisions for heat removal.

Due to space constraints, the magnet can be no larger than 6" in diameter. The thickness must be kept to a minimum in order to keep the overall height of the transport module (mounted on top of this magnet) and its center of gravity low. It will be made out of mild steel (0.2% carbon) which is a typical machining steel and readily available. As required

by the design of the kinematic coupling, the two halves will be separated from one another by a small gap. The size of this gap will be specified through the course of designing the magnet.

6.1 Theory of operation

The force that an electromagnet provides is dependant on a number of variables. Calculating this force requires the examination of the shape of the magnet and the properties of the material used to construct the magnet. The following derivation of the force is based on a similar discussion in *Electromechanical Dynamics* by Woodson and Melcher [18].

First, the mechanisms for mechanical work in a magnetoquasistatic system need to be considered from a power balance perspective. Doing so will provide the insight as to how the force is generated by an electromagnet. Ultimately, an expression for the force in terms of the electrical inputs and variables will be found. Assume that the system is lossless. The losses in the system (resistance, damping, etc.) can be accounted for outside of the system.

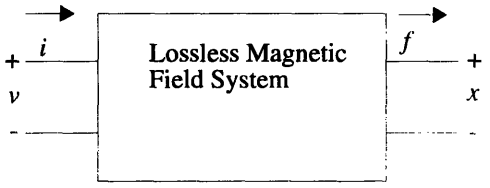


Figure 6.2: Four terminal representation of a magnetic field system.

The electromagnet can be represented as a two-port system. On one side are the electrical variables, voltage v and current i. On the other side are the mechanical variables, force f and displacement x [18].

The power balance equation for the system [8] with terminal variables as shown in Figure 6.2 is

$$vi = \frac{dw}{dt} + f\frac{dx}{dt} \tag{6.1}$$

where v and i are voltage and current inputs, w is the stored energy in the magnetic circuit, f and x are force and displacement. Using Faraday's law [8] for a coil and after some rearrangement, (6.1) can be expressed as

$$\frac{dw}{dt} = i\frac{d\lambda}{dt} - f\frac{dx}{dt} \tag{6.2}$$

where λ is the flux linkage, and λ and i are independent variables. If the energy w is to be conserved in the system, the energy, as a function of λ and x, must be single-valued with finite second partial derivatives with respect to λ and x. Thus, the first partial derivatives must also be single-valued and small changes in energy can be expressed as

$$dw = \frac{dw}{d\lambda}\lambda + \frac{dw}{dx}x. \tag{6.3}$$

Based on (6.2) and (6.3), the following two relationships can be derived:

$$i = \frac{dw}{d\lambda} \tag{6.4}$$

$$f = -\frac{dw}{dx}. \tag{6.5}$$

Ultimately, this mechanical force, f, produced by the system will be determined.

It is necessary to determine an expression for the energy stored, W, in the system. After multiplying both sides of (6.2) by dt, this equation becomes

$$dw = id\lambda - fdx. \tag{6.6}$$

This equation describes the change in energy for small changes in λ and/or x. For the system described in Figure 6.2, if λ and x are chosen as independent variables, it follows that the current and the force must be functions of these two variables [18]. That is, i and f can be written as

$$i = i(\lambda, x) \tag{6.7}$$

$$f = f(\lambda, x) \ . \tag{6.8}$$

Using (6.7) and (6.8), (6.6) can be rewritten as a function of only two variables, λ and x:

$$dw = i(\lambda, x)\, d\lambda - f(\lambda, x)\, dx. \tag{6.9}$$

This equation describes the change in energy given a change in the independent variables, λ and x. To find the total stored energy, the change in energy must be evaluated over the amounts that λ and x changes. In the electromagnetic system considered here, assume that everything starts from rest, $\lambda = 0$ and $x = 0$, and increases to $\lambda = \bar{\lambda}$ and $x = \bar{x}$. Thus, (6.9) must be integrated over $\lambda \in (0, \bar{\lambda})$ and $x \in (0, \bar{x})$. In terms of the variable space where λ and x are independent (orthogonal) variables shown in Figure 6.3, (6.9) must be

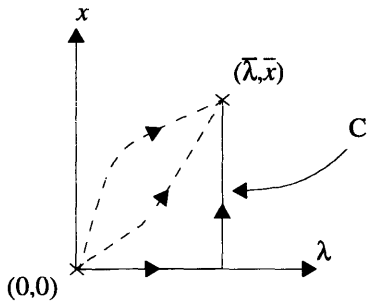


Figure 6.3: Variable space of λ and x

integrated from the point $(0, 0)$ to the point $(\bar{\lambda}, \bar{x})$. There are an infinite number of possible paths of integration, three of which are shown in Figure 6.3 [18]. But since energy is conserved in this lossless system, the energy is independent of the path of integration from point $(0, 0)$ to the point $(\bar{\lambda}, \bar{x})$. Thus, any path can be used. The path C shown in Figure 6.3 is a good choice since it breaks the integration into two integrations each over a single variable. Along this path, x first changes while $\lambda = 0$, then λ changes while $x = \bar{x}$. Thus, the integral of (6.9), which gives an expression for the stored energy W, takes on the form

$$W(\bar{\lambda}, \bar{x}) - W(0, 0) = -\int_0^{\bar{x}} f(0, x)\, dx + \int_0^{\bar{\lambda}} i(\lambda, \bar{x})\, d\lambda . \tag{6.10}$$

The energy at $\lambda = 0$ and $x = 0$ is zero since it was assumed that everything starts at rest. Also, the force when $\lambda = 0$ is zero as well. With these design assumptions, (6.10) reduces to

$$W(\bar{\lambda}, \bar{x}) = \int_0^{\bar{\lambda}} i(\lambda, \bar{x})\, d\lambda \tag{6.11}$$

which describes the stored energy as a function of the magnetic flux linkage λ at some fixed displacement $\bar{x}$.

From [8], the flux linkage and the current can be related by the equation

$$\lambda = \mathrm{L}(x)\, i \tag{6.12}$$

where the $\mathrm{L}(x)$ is the inductance as a function of the displacement x.

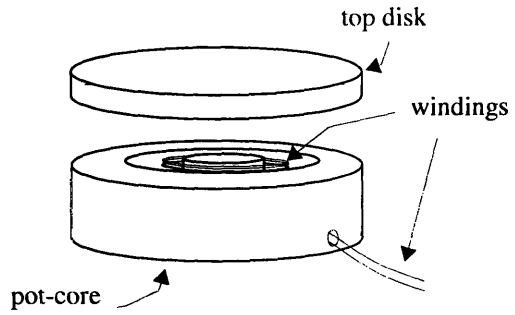


Figure 6.4: Geometry of electromagnet

Equation (6.11) can consequently be rewritten as

$$W(\bar{\lambda}, \bar{x}) = \int_0^{\bar{\lambda}} \frac{\lambda}{\mathrm{L}(x)} d\lambda = \frac{\lambda^2}{2\mathrm{L}(x)}. \tag{6.13}$$

Given $W(\bar{\lambda},\bar{x})$ in (6.13), the force can be determined from (6.5) to be:

$$\begin{aligned} f &= -\left.\frac{dw}{dx}\right|_{\lambda} \\ &= \frac{-\lambda^2}{2L(x)} \cdot \frac{d}{dx}\mathrm{L}(x) \\ &= \frac{1}{2} i^2 \frac{d}{dx}\mathrm{L}(x) \end{aligned} \tag{6.14}$$

This expression for the force is dependant on the magnitude of the current and not its direction. The force also acts in the direction of positive x, i.e. in the direction that increases the inductance of the system. In order to determine the force that the magnet can exert, an expression for the inductance must be found. However, in order to do so, the material properties of the electromagnetic circuit and its basic geometry must be known.

6.2 Preliminary Analysis

The basic geometry of the electromagnet is shown in Figure 6.4. It is a circular pot-core type design with the current windings wrapped around the center yoke. There is an air gap between the pot-core and a top disk that completes the magnetic circuit. If the force and the magnetizing current desired are known, then the first attempt at designing the magnet should answer two questions. First, how many turns of wire around the center yoke are required? Second, how does the size of the air gap impact this number? The answers to these questions give a rough idea of the feasibility of using an electromagnet

for this clamping application. For example, if the number of turns required for a given magnetizing current exceeds the number which can be placed in the available winding space, it may not be feasible to use an electromagnet to provide the desired force. Similarly, if the air gap must be smaller than the tolerance that can achieved in machining the parts, then an electromagnet may not be the right clamp for this application.

To simplify calculations in this preliminary analysis, make the following two assumptions:

1. The iron is infinitely permeable, $\mu = \infty$.
2. The magnetic flux is contained within the iron and takes the shortest path across the air gap.

The first assumption is a reasonable one for a first approximation since the permeability of unsaturated ferromagnetic materials are two to three orders of magnitude greater than the permeability of free space. The second assumption follows from the first; the flux will tend to flow in the most permeable substance and consequently, prefer the iron over free space.

The first assumption implies that the magnetic field in the iron is zero. With $\mathbf{B} = \mu\mathbf{H}$ and the condition that the flux density is continuous across boundaries of different permeability [8], the following condition results:

$$\mu_0 \mathbf{H}_{gap} = \mu \mathbf{H}_{iron}, \tag{6.15}$$

where μ_0 is the permeability of free space ($4\pi \times 10^{-7}$ [Henry/meter]). When a magnetic field is applied in the gap, the left hand side of (6.15) is some finite flux density. However, the right hand side must also be that same finite number. Thus, the infinite permeability assumption implies that the magnetic field in the material is effectively zero. This assumption will be reconsidered later.

To determine the inductance, Ampere's (6.16) and Gauss's (6.17) Laws must be solved simultaneously to arrive at a relationship between the flux linkage and the current [10].

$$\oint \mathbf{H} \cdot d\mathbf{l} = \int \mathbf{J} \cdot d\mathbf{l} \tag{6.16}$$

$$\int \mathbf{B} \cdot d\mathbf{S} = 0 \tag{6.17}$$

Since **H** in the iron is zero, Ampere's Law simplifies to

$$\mathbf{H}_1 x + \mathbf{H}_3 x = Ni \tag{6.18}$$

where $\mathbf{H}_1$ and $\mathbf{H}_3$ are the field intensities in the gap x, N is the number of turns, and i is the magnetizing current. As shown in Figure 6.5, $\mathbf{H}_1$ is the field in the gap between the core and the top plate while $\mathbf{H}_3$ is the field in the outer ring of the magnet.

Using Gauss's Law and defining a closed surface as shown in Figure 6.5, the following relations can be found:

$$\mu_0 \mathbf{H}_1 (A_1) - \mu_0 \mathbf{H}_3 (A_2) = 0$$

$$\mathbf{H}_3 = \frac{A_1}{A_2} (\mathbf{H}_1) \,. \tag{6.19}$$

Combining (6.18) and (6.19), the magnetic field $\mathbf{H}_1$ is:

$$\mathbf{H}_1 = \frac{Ni}{\left(1 + \frac{A_1}{A_2}\right) x} \tag{6.20}$$

Note that there is no dependency on the geometry of the magnet with the exception of the two cross-sectional areas and the air gap between the halves of the coupling. This is the result of the infinite permeability assumption.

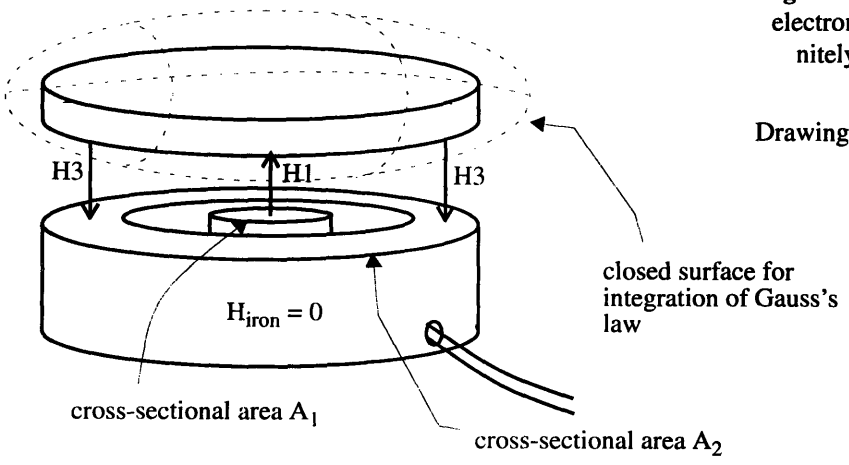


Figure 6.5: Magnetic fields in an electromagnet made of an infinitely permeable material.

Drawing is not to scale

Since the magnetic flux is defined as the flux density times the area, it can be expressed as

$$\Phi = \mathbf{B}A_1 = \mu_0 \mathbf{H}_1 A_1 \tag{6.21}$$

The cross-sectional area used is that of the core since by assumption 2, the extent of the shortest path taken across the air gap is bounded by the area of the core of the magnet, A_1. The flux linked by N turns of wire is

$$\lambda = N\Phi = \frac{\mu_0 N^2 A_1 i}{\left(1 + \frac{A_1}{A_2}\right)x}. \tag{6.22}$$

By comparing (6.22) with (6.12), the inductance and its first derivative with respect to the gap x is:

$$\mathrm{L}(x) = \frac{\mu_0 N^2 A_1}{\left(1 + \frac{A_1}{A_2}\right)x}$$

$$\frac{d\mathrm{L}}{dx} = -\frac{\mu_0 N^2 A_1}{\left(1 + \frac{A_1}{A_2}\right)x^2}. \tag{6.23}$$

Finally, by substituting (6.23) into (6.14), the force is

$$f = \frac{1}{2} i^2 \left(\frac{\mu_0 N^2 A_1}{\left(1 + \frac{A_1}{A_2}\right)x^2} \right). \tag{6.24}$$

This is an expression for the upper limit of the force. The assumption of infinite permeability made any effects due to the geometry, other than the two cross-sectional areas and the air gap, negligible. As shown in chapter 6.4, the effects of the shape of the magnet serves to decrease the effective force when taken into account. It does, however, make explicit the important fact that the force is inversely proportional to the square of the air gap.

Equation (6.24) is also good for a feasibility check. For this clamping application, a one Ampere current was chosen to provide a force of 150 pounds - a 50% margin beyond

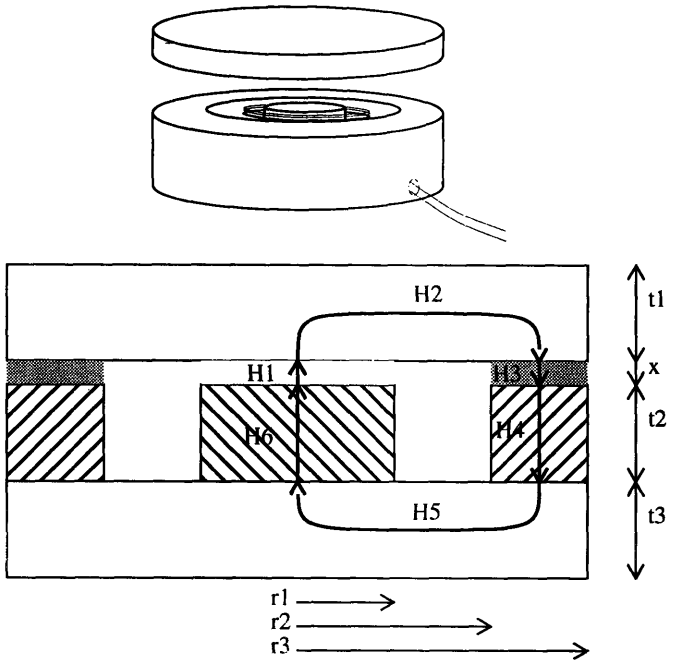


Figure 6.6: Dimensions of the electromagnet

The dimensions are labelled below. The magnetic fields is 6 sections are also labelled.

the 100 pound specification. Since the electromagnet could be no larger than 6", reasonable choices for the radius of the core and the width of the outer ring was 1" each. This provided a 1" gap for the windings. A_1 and A_2 are respectively 3.14 square inches and 15.71 square inches. The depth of the magnet as well as the thickness of the top and bottom disks were chosen to be roughly 1". Figure 6.6 labels these dimensions on a cross-sectional drawing of magnet. After some trial and error on a spread sheet program [11], a reasonable number of turns and gap size were 200 turns and 0.010" respectively.

6.3 Determination of the material permeability

The first attempt was hardly enough to be considered a complete treatment of the force of the magnet, though it was sufficient to warrant pursuing the idea for the kinematic coupling application further. The assumption of an infinitely permeable material allowed us to neglect most of the geometry of the magnet and simplify the calculation at the cost of accuracy. Before attempting a deeper analysis of the magnetic circuit of Figure 6.4, the permeability of the iron μ must be determined, or at least well estimated since it determines how much of an effect the overall shape of the magnet has on the force it produces.

The permeability μ of ferromagnetic materials is much greater than that of free space which consequently makes it a very good "conductor" of magnetic flux. However, it is not infinitely permeable and will saturate given a high enough applied magnetic field. Since $\mathbf{B} = \mu\mathbf{H}$, in order to reflect the saturation limit of a ferromagnetic specimen, the permeability must be a non-linear function of the field **H**. Typically, the permeability is found as a result of experimentally measuring the magnetic flux given a range of applied fields. The permeability is the slope of the resulting curve. Since such experiments are well-documented, there is no need to perform them to determine μ for the material used. Rather, it is sufficient to find the curve for the material and make the best estimate of μ possible from it. Figure 6.7 shows the **H** versus **B** characteristic for a typical commercial low-carbon iron.

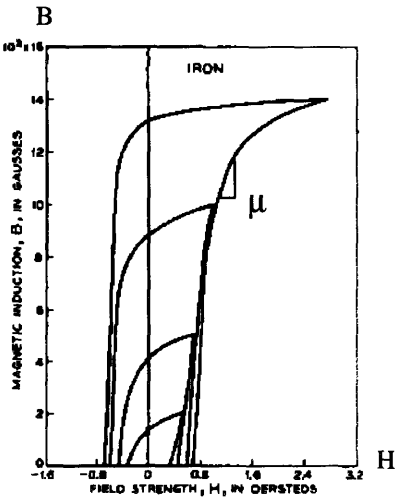


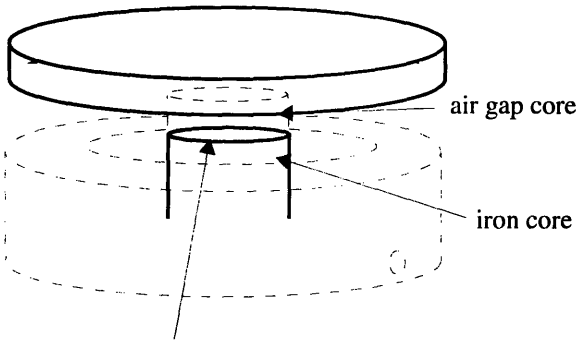
Figure 6.7: Typical magnetic field versus magnetic flux characteristic of iron.

The permeability μ is the slope of this curve (as shown). This graph was copied from *Ferromagnetism* by Bozorth [3].

First, a data set in Matlab (found in the "Flexible Assembly Project Software Technical Report" [11]) can be created by using sample points of the original B-H curve. By low-pass interpolating the sample points and then oversampling this curve, Matlab can construct a data set of two finely-spaced vectors corresponding to values of B and H. The permeability at almost any point of the curve can be found by taking the ratio of the difference of two adjacent pairs of B and H, i.e.

$$\mu = \frac{B_n - B_{n-1}}{H_n - H_{n-1}}$$

Figure 6.8: Illustration of regions in the core of the electromagnet



cross-sectional area, A_1, of iron core = cross-sectional area of air gap

The resolution of the values of the permeability depends on how many times the original sample set was oversampled after low-pass interpolation.

Given a data set with the B-H curve for the material being used and the applied magnetizing current, the most straightforward method of determining the permeability is to search for it iteratively utilizing this B-H data set. Essentially, starting from an initial guess of the permeability of the iron μ_{guess}, determine the flux density **B** in the iron. Using the B-H data set, estimate the permeability $\mu_{estimate}$ associated with this value of **B**, and compare it to the original guess. If the original guess equals this estimate, the permeability for the given magnetizing current has been found. Otherwise, the guess should be modified to reflect the difference between the guessed and estimated permeability and the previous steps must be performed again. These steps can be performed iteratively until the difference between the guessed and the estimated permeabilities diminishes to an acceptable value.

The difficulty in implementing the above procedure lies in the determination of the magnetic flux density in different sections of the iron. The flux density must be calculated for every section having different cross-sectional area in order to determine the permeability in each of those sections. The flux through the circuit is constant for a given applied field. However, since the cross-sectional areas are different at different points in the circuit, the flux density differs as well. Consequently, the permeability, which is a function of the density, is also different in every region. For regions of free space and iron where the

cross-sectional areas are the same, these densities, namely $\mathbf{B_1}$, $\mathbf{B_6}$ and $\mathbf{B_3}$, $\mathbf{B_4}$, can best be found by first evaluating the magnetic field in the regions of free space, $\mathbf{H_1}$ and $\mathbf{H_3}$. Refer to Figure 6.6. Since the cross-sectional areas are the same, the density in the air gap equals the density in the iron. That is,

$$\mathbf{B_6} = \mathbf{B_1} = \mu_0 \mathbf{H_1} \tag{6.25}$$

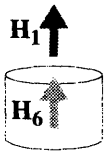
and

$$\mathbf{B_4} = \mathbf{B_3} = \mu_0 \mathbf{H_3} \tag{6.26}$$

Having found $\mathbf{B_6}$ and $\mathbf{B_4}$, the permeabilities, μ_6 and μ_4, can be found from the B-H data set in Matlab. Again, the differences between the guessed and estimate permeabilities can be used to modify the guesses until their differences are small. Thus, if the magnetic fields in the air gaps, $\mathbf{H_1}$ and $\mathbf{H_3}$, can be found, then the permeability of the iron, μ_6 and μ_4, (in sections 6 and 4, respectively) can be estimated. The permeabilities, μ_2 and μ_5, will later be estimated from μ_6 and μ_4.

To find $\mathbf{H_1}$ and $\mathbf{H_3}$, use Gauss's Law to describe the magnetic field in each section of the magnet in terms of the field in an air gap. Then application of Ampere's Law yields the fields in terms of the *N*, *i*, and the geometry of the magnet. Figure 6.9 describes this process for calculation of $\mathbf{H_1}$. The form that $\mathbf{H_1}$ takes is

$$\mathbf{H_1} = \frac{Ni}{\left(1+\frac{A_1}{A_2}\right)x + \frac{\mu_0}{\mu}\left(2l + \left(1+\frac{A_1}{A_2}\right)t_2\right)}. \tag{6.27}$$



$$\mu_0 H_1 A_1 = \mu H_6 A_1$$

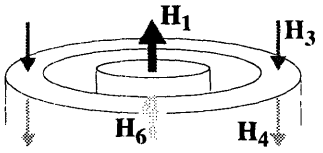
$$\frac{\mu_0 H_1}{\mu} = H_6$$

Figure 6.9: Calculation of $\mathbf{H_1}$ in terms of all other $\mathbf{H_j}$.

Gauss's Law makes it possible to write $\mathbf{H}_j$ for j between 2 and 6 in terms of $\mathbf{H_1}$.

$$\int \mathbf{B} \cdot d\mathbf{S} = 0$$

Closed surfaces for integration are implied in figures on left. Dimensions for magnet can be found in Figure 6.6.



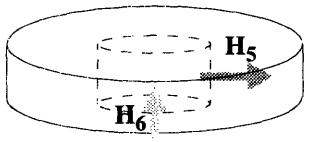
$$\mu_0 H_1 A_1 = \mu H_3 A_2$$

$$\frac{A_1}{A_2} H_1 = H_3$$

$$\mu_0 H_6 A_1 = \mu H_4 A_2$$

$$\frac{A_1}{A_2} H_6 = H_4$$

$$\frac{A_1}{A_2} \cdot \frac{\mu_0 H_1}{\mu} = H_4$$



$$\mu H_6 (2\pi r_1 t_1) = \mu H_5 \left[2\pi t_1 \left(\frac{r_3 + r_2}{2} \right) \right]$$

$$\frac{r_1}{\frac{r_3 + r_2}{2}} \cdot \frac{\mu_0 H_1}{\mu} = H_5 = H_2$$

After solving for $\mathbf{H_1}$ in terms of all other $\mathbf{H}$, write Ampere's Law in terms of $\mathbf{H_1}$, N, and i

$$\sum_{j=1}^{6} \mathbf{H}_j l_j = Ni$$

$$\mathbf{H}_1 = \frac{Ni}{\left(1 + \frac{A_1}{A_2}\right)x + \frac{\mu_0}{\mu}\left(2l + \left(1 + \frac{A_1}{A_2}\right)t_2\right)}$$

Following a similar process, $\mathbf{H_3}$ can be found to be

$$\mathbf{H}_3 = \frac{Ni}{\left(1 + \frac{A_2}{A_1}\right)x + \frac{\mu_0}{\mu}\left(2l + \left(1 + \frac{A_2}{A_1}\right)t_2\right)}. \tag{6.28}$$

A Matlab script in [11] was created to carry out the estimation of μ_6 and μ_4, given (6.27) and (6.28) with the values of the variables and a data set containing values of B and

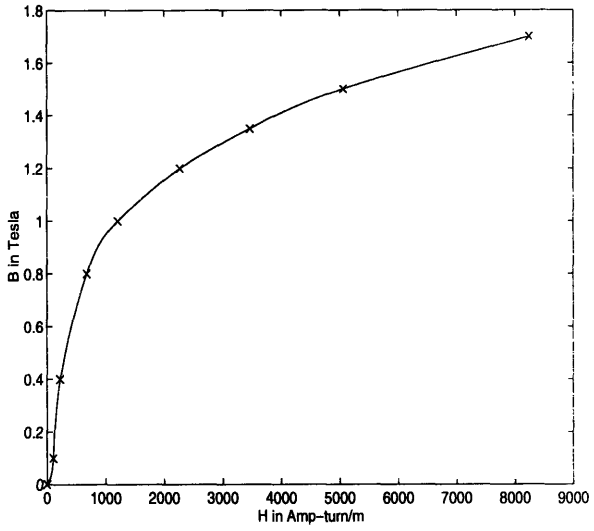


Figure 6.10: B-H curve created for the calculation of μ_6 and μ_4.

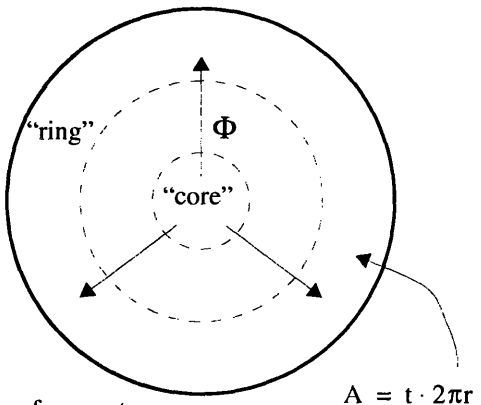
The Matlab that created this curve can be found in Appendix F.

H. Figure 6.10 contains a plot of the B-H curve that was used to estimate μ_6 and μ_4. The results of the calculation are shown in Table 6.2.

current [Amperes]	μ_6 [Hm^{-1}]	relative μ_6 (μ_6/μ_0)	μ_4 [Hm^{-1}]	relative μ_4 (μ_4/μ_0)
0.50	2.1436e-3	1705.8	7.1359e-4	567.9
0.75	1.5270e-3	1215.1	7.9896e-4	635.8
1.00	1.2694e-3	1010.2	9.3108e-4	741.9
1.25	1.0879e-3	865.7	1.4374e-3	1143.9
1.50	9.0235e-4	718.1	2.5131e-3	1999.8
1.75	7.9592e-4	633.4	4.8630e-3	3869.8
2.00	6.6566e-4	529.7	3.6124e-3	2874.6
2.25	5.5550e-4	442.1	2.8274e-3	2250.0
2.50	5.2341e-4	416.5	2.3696e-3	1885.7
2.75	4.2032e-4	334.5	2.1048e-3	1674.9
3.00	3.8513e-4	306.5	1.8918e-3	1505.4

Table 6.1: Permeabilities, μ_6 and μ_4, given input current

For the other permeabilities, μ_2 and μ_5, finding an expression for a magnetic field in free space with the same cross-sectional area as the top and bottom disks of the magnet (section 2 and 5 in Figure 6.6) is not possible. An additional complication to determining these permeabilities is that the cross-sectional area changes as the flux flows from the cen-



Top view of magnet
Dashed lines indicate boundary of "core" and "ring".

$A = t \cdot 2\pi r$ where $r \in (1, 2.25)$

Figure 6.11: Change in permeability with cross-sectional area.

As Φ radiates outward (or inward) the cross-sectional area, A, changes. Consequently, the density **B** and the permeability μ changes.

ter of the "disk" to the circumference as shown in Figure 6.11. The cross-sectional area changes from a minimum of 0 square inches to 18.8 square inches at the circumference of the magnet. However, since the two sections are essentially identical disks, the permeabilities will be equal. A convenient approximation for μ_2 and μ_5 is to take the average of μ_6 and μ_4. It is convenient because μ_6 and μ_4 have already been determined in the preceding paragraphs. μ_6 and μ_4 also correspond to cross-sectional areas that are within the range of areas of the disk, 0 to 18.8 square inches. μ_6 was the permeability of the core with a cross-sectional area of 3.14 square inches, while μ_4 was the permeability of the ring with a cross-sectional area of 15.7 square inches. Thus, at some point on the disk, μ_2 (and equivalently μ_5) will take on the values of μ_6 and μ_4. Taking the average of μ_6 and μ_4 will be close to determining the actual average permeability μ_2.

The estimated values of μ_2 and μ_5 are calculated from the values of μ_6 and μ_4 in Table 6.2 and listed in Table 6.2.

current [Amperes]	μ_2, μ_5 [Hm^{-1}]	relative permeability (μ_2/μ_0)
0.50	1.4286e-3	1137.8
0.75	1.1630e-3	925.5
1.00	1.1003e-3	875.6
1.25	1.2627e-3	1004.8
1.50	1.7077e-3	1358.9
1.75	2.8294e-3	2251.6
2.00	2.1390e-3	1702.2
2.25	1.6915e-3	1346.1
2.50	1.4465e-3	1151.1
2.75	1.2626e-3	1004.7
3.00	1.1385e-3	906.0

Table 6.2: Permeabilities, μ_2 and μ_5, given input current

6.4 Improved Analysis

With the knowledge of the values of the permeabilities in each section of the magnet, it is possible to determine a better estimate of the force this electromagnet can provide.

In evaluating Ampere's law once again, it is useful to approximate the line integral by a summation of sections, each taken over a length having a constant value of μ and a definable cross-sectional area

$$\oint \mathbf{H} \cdot d\mathbf{l} \cong \sum_{i}^{6} H_i l_i = Ni \quad . \tag{6.29}$$

The division into sections is shown in Figure 6.6. Equation (6.17), Gauss's law, implies that the lines of **B** are continuous around the magnetic circuit. Therefore, the flux Φ pass-

ing through any cross-section of the circuit is the same. The value of **H** in any section can be related to Φ by

$$\Phi = \mathbf{B}_j A_j = \mu_j \mathbf{H}_j A_j \tag{6.30}$$

where the μ_j and A_j, the permeability of the metal and the cross-sectional area in division *j*, must be taken for that section. Thus, the magnetic field in each section can be expressed as

$$\mathbf{H}_j = \frac{\Phi}{\mu_j A_j}. \tag{6.31}$$

Substitution of (6.31) into equation (6.29) and re-arranging the variables yields

$$Ni = \Phi \sum_{j=1}^{6} \frac{l_j}{\mu_j A_j} = \Phi R \tag{6.32}$$

which relates the magnetizing current to the flux. The term, $\frac{l_j}{\mu_j A_j}$, is known as the reluctance R_j for each section [10].

Knowing the value of the total reluctance is useful in determining the inductance of the circuit so it will be calculated here first before continuing. Equation (6.32) may be thought of as a circuit analogous to a series electric circuit of the form $V = I \cdot \left(\sum_j R_j \right)$. In this magnetic circuit, the driving or magnetomotive force *Ni*, the flux Φ, and the reluctances R_j are analogous to the voltage V, the current I, and the resistances R_j of the electrical circuit, respectively [10]. This analogy is shown in Figure 6.12. Consequently, to solve for the total reluctance in this series magnetic circuit, calculate the reluctance for each division *j* and add them together:

$$\begin{aligned} R &= R_1 + R_2 + R_3 + R_4 + R_5 + R_6 \\ &= \frac{x}{\mu_0 A_1} + \frac{t_2}{\mu A_1} + \frac{2}{\mu 2\pi t_1} + \frac{x}{\mu_0 A_2} + \frac{t_2}{\mu A_2} \end{aligned} \tag{6.33}$$

Figure 6.12 shows how the reluctances for each section are calculated using the dimensions specified in Figure 6.6.

$$R_1 = \frac{x}{\mu_0\left(\pi r_1^2\right)} = \frac{x}{\mu_0 A_1}$$

$$R_2 = \frac{1}{\mu 2\pi t_1} = R_5$$

$$R_3 = \frac{x}{\mu_0\pi\left(r_3^2 - r_2^2\right)} = \frac{x}{\mu_0 A_2}$$

$$R_4 = \frac{t_2}{\mu\pi\left(r_3^2 - r_2^2\right)} = \frac{t_2}{\mu A_2}$$

$$R_6 = \frac{t_2}{\mu\left(\pi r_1^2\right)} = \frac{t_2}{\mu A_1}$$

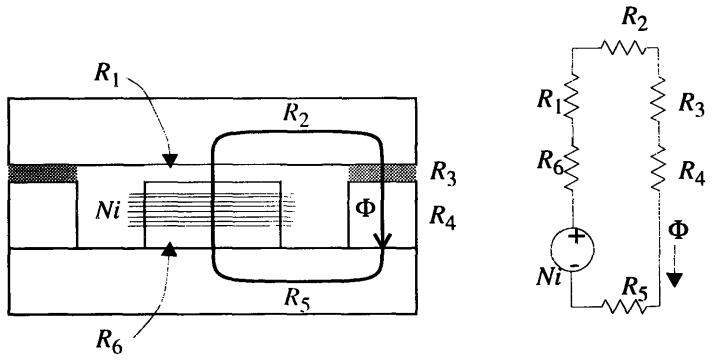


Figure 6.12: Reluctance calculations
The electromagnet with the reluctance of each section is shown as well as the electrical circuit analog for a pot-core electromagnet.

Given that the total reluctance is known, the inductance of the electromagnetic circuit can be found. Since the flux linkage λ is the flux Φ multiplied by the number of turns of the coil N, multiplying both sides of (6.32) by N and rearranging terms gives

$$\lambda = \frac{N^2}{\sum_{j=1}^{6} \frac{l_j}{\mu_j A_j}} i = \frac{N^2}{R} i. \tag{6.34}$$

Thus, comparing (6.12) with (6.34), the inductance due to the air gap must be

$$\mathrm{L}(x) = \frac{N^2}{\sum_{j=1}^{6} \frac{l_j}{\mu_j A_j}} = \frac{N^2}{R}. \tag{6.35}$$

Substituting (6.33) into (6.35), the inductance and its derivative with respect to x becomes

$$\mathrm{L}(x) = \frac{N^2}{R} = \frac{N^2}{\frac{x}{\mu_0 A_1} + \frac{t_2}{\mu A_1} + \frac{2}{\mu 2\pi t_1} + \frac{x}{\mu_0 A_2} + \frac{t_2}{\mu A_2}},$$

$$\frac{d}{dx}\mathrm{L}(x) = \frac{-N^2}{\mu_0 R^2}\left(\frac{A_2 + A_1}{A_1 A_2}\right). \tag{6.36}$$

Substituting (6.36) into (6.14), the force that the magnet is expected to exert is

$$f = \frac{1}{2} i^2 \left[\frac{N^2}{\mu_0 R^2} \left(\frac{A_2 + A_1}{A_1 A_2} \right) \right]. \tag{6.37}$$

This is the force that this magnet should exert given the magnetizing current, the number of turns of the coil, the permeability of the iron used, and the geometry.

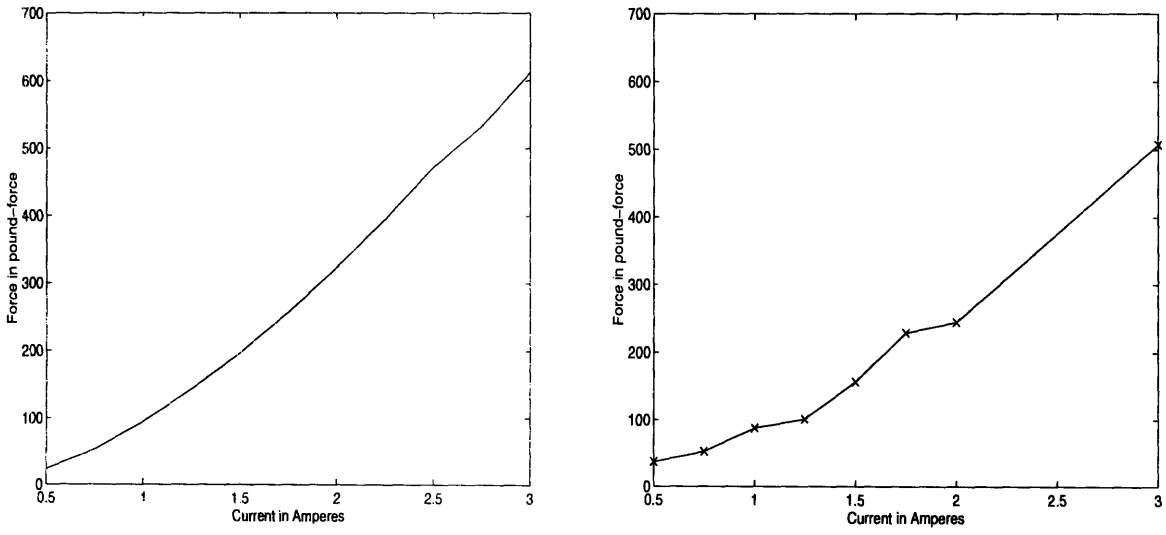
For the dimensions specified in Figure 6.6 and utilizing the estimated values of the permeability in Table 6.2, the force was estimated using (6.37) for magnetizing currents between 0.5 and 3 Amperes. The values are listed in Table 6.3 and plotted in Figure 6.13.

current in Amperes	force in lbf
0.50	24.1
0.75	53.2
1.00	94.6
1.25	144.3
1.50	196.9
1.75	256.7
2.00	323.8
2.25	395.3

Table 6.3: Theoretical results of maximum clamping force given a magnetizing current

Figure 6.13: Plots of force versus current

The left hand plot is a theoretical plot; it was created by estimating the force given the geometry of the magnet, the permeabilities of the iron, and the current. The right hand plot is the results of a set of tests of the magnet's clamping force using an Instron machine.



current in Amperes	force in lbf
2.50	472.1
2.75	530.0
3.00	613.0

Table 6.3: Theoretical results of maximum clamping force given a magnetizing current

6.5 Experimental results

The actual dimensions of the magnet are included in Figure 6.6. Using an Instron machine at the Mechanical Behavior of Materials Laboratory[1] at the Massachusetts Institute of Technology, a set of experiments were carried out to determine the clamping capabilities of the magnet. With the magnet driven by a fixed DC current, the Instron machine slowly applied force at a rate of 10 lbf/second to pull the halves of the magnet apart. The force which caused the magnet to fail was recorded and plotted in Figure 6.13. Table 6.4

current in Amperes	force in lbf
0.50	37.34
0.75	52.7
1.00	87.9
1.25	101.0
1.50	157.4
1.75	228.4
2.00	243.8
3.00	507.4

Table 6.4: Experimental data of maximum clamping force given a current

lists the forces when the magnet failed for a given current. Both theoretical and experimental curves shows that the force is not just a function of the current; otherwise, it would have grown quadratically. Rather, the drop in the permeability at higher flux densities, corresponding to the saturation of the material, attenuates the amount of force that the magnet

1. These experiments were carried out with the assistance of Professor Mohammed Durali.

is able to supply given a current. The most likely source of error between the theoretical and experimental plots in Figure 6.13 arises from the fact that the actual B-H curve of the electromagnet was not verified. Thus, the permeabilities that were estimated in chapter 6.3 may be slightly inaccurate. Another source of error is possibly due to an assumption that was made concerning the losses in the system. It was assumed that the lines of flux were confined to the material of the circuit. This is not exactly true and thus, could give rise to the error in the plots.

7 Sensing for Product Location

Manufacturing products with small tolerances often require that the assembly machines be built with very tight tolerances. This is because the machine becomes the frame of reference in which to locate the product casing and features on the product. A product casing is generally located against some surface on the assembly machine whose position is known very accurately; this surface, thus, becomes a frame of reference. Specific features on the casing are then found by inferring its location from this surface. This implies that the manufacturing tolerance of the surface used on the part also be made accurately, whether or not it is significant to the end function or purpose of the casing. It also implies that the relationship of the surface used and the actual feature of interest be known exactly and be consistent for all products of the same type. For example, on a product casing, the feature of interest may be a hole whose position is calculated from knowing the location of the side of the casing and the relationship between that side and the hole. The repeatability of locating this hole heavily depends on maintaining tight geometric tolerances on the casing and the assembly machine. This makes manufacturing casings costly. Furthermore, as the demand for smaller and smaller tolerances exceeds what can be accomplished by conventional mechanical means (such as using a mating surface as described above), alternate methods must be found that can directly locate the features of interest rather than inferring their locations. In order for these methods to be economically viable, they must locate features as well as, if not better than by conventional means, and they must also be less expensive to implement.

This section documents preliminary studies of one alternate method: using a light source, photodetector, and a small-diameter fiber-optic strand to locate the small holes in a product casing. The typical product casing is made of solid molded plastic with regularly-spaced openings for the parts that are inserted into it. These openings run throughout the width of the casing. Figure 7.1 is a picture of a typical product[1]. The problem is to locate these holes directly without relying on some other feature on the product. It is assumed that the approximate location of the hole is known based on the proper placement of the

1. Figure 7.1 was prepared by Miguel Barrientos [2].

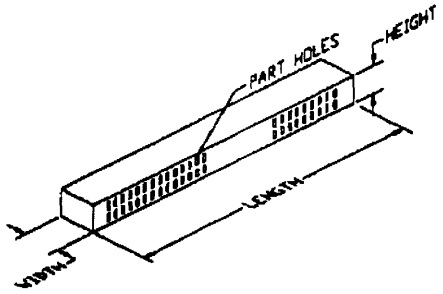


Figure 7.1: Typical product casing

product with respect to the photodetector. The role of this sensor is then to determine the exact location of the hole. It is also assumed that if the product was backlit, the light which passes comes only through the holes - the features that need to be measured. The concept behind using a photodetector and light source as a backlight is to try to determine the location of the holes by using the silhouette of the backlit product.

7.1 Experimental Setup

The setup for the following experiments consisted of five components: a phototransistor circuit, a fiber-optic probe and housing for the photo transistor, an x-z linear stage with micrometer adjustments, a 500 Watt halogen bulb, and an oscilloscope.

The phototransistor circuit is shown in Figure 7.2. It basically consists of a phototrans-

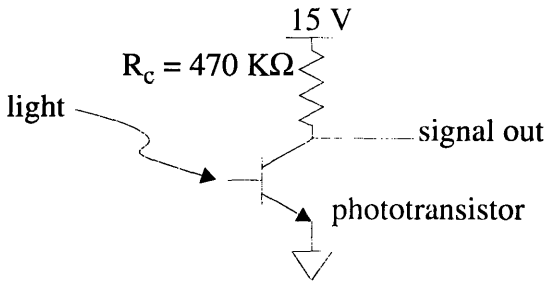


Figure 7.2: Phototransistor circuit.

istor hooked up in a common collector configuration with a collector resistor, $R_c = 470\text{K}\Omega$ connected to a 15 volt DC power supply. The large resistance was necessary because the light intensities that the phototransistor received were very small. Thus, the input needed to be amplified; in this case, for a transistor with a gain of approximately $\beta = 100$, the transconductance gain was $\beta R_c = 4.7\times10^8$.

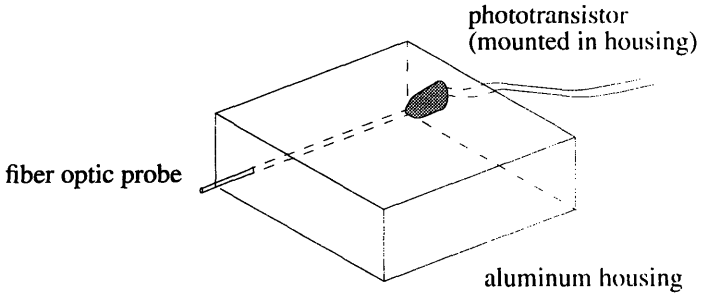


Figure 7.3: Aluminum housing for phototransistor and fiber optic probe.

The phototransistor was mounted in an aluminum housing, shown in Figure 7.3, and shielded from all light except that which came through a 0.022" opening. A 2" length of a 0.020" diameter plastic fiber-optic strand was sheathed in a hypodermic steel tubing and inserted half way into the aluminum housing. The fiber acted as a light pipe, transmitting only the light that shines directly into it to pass to the phototransistor. Thus, the fiber optic strand and sensor essentially functioned as a probe with a 0.020" diameter aperture.

The product itself was mounted on an x-z linear stage with micrometer adjustments. That is, the product can be moved horizontally and vertically with respect to the fiber probe with an accuracy of 0.001".

The light source used for these experiments was a 4" by 0.375" 500 Watt halogen lamp located roughly 3" away from the product. The light output was controlled by a dimmer. Since the light source was the same size and shape as the product, the light which came from it through the product was assumed to be uniform.

Finally, a digital oscilloscope was used to measure both the perceived and actual light intensities. The perceived light intensities was measured from the collector of the phototransistor. The light output of the lamp was measured indirectly by monitoring the current drawn by the light. A Hall effect sensor attached to the power cable of the lamp produced a voltage that was proportional to the current drawn by the lamp.

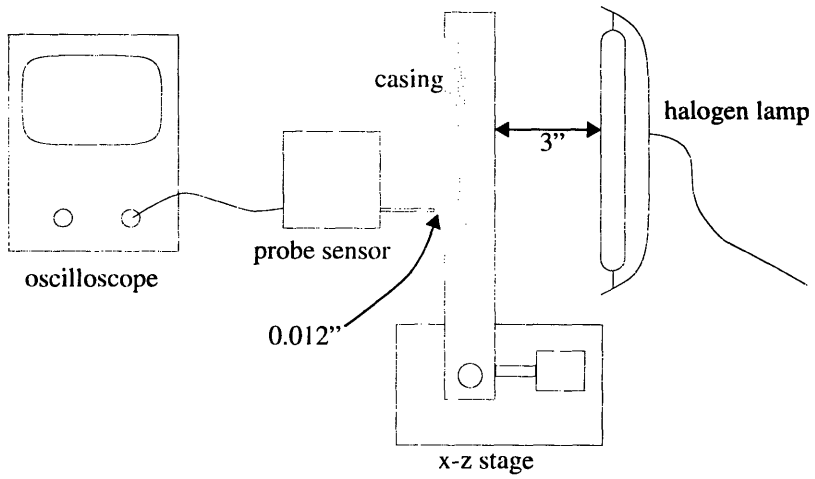


Figure 7.4: Experimental setup

A top view of the complete setup is shown in Figure 7.4. The tip of the probe was placed 0.012" from the surface of the casing to minimize the sensitivity of the photodetector to stray light in the environment. The alignment target that was used in the experiments is shown in Figure 7.1. The light source was mounted roughly 3" behind and parallel to the product, as shown. The intense heat from the lamp prevented closer placement of the light.

7.2 Experiments

Three sets experiments were conducted - each set at a different light output from the lamp. For each setting, several measurements of the position were made in order to locate the upper left hole along the horizontal x-axis - the axis parallel with the target - and along the vertical z-axis. That is, the target was swept across the field of view of the sensor first along the horizontal axis, then second along the vertical axis. During movements along one axis, the position was recorded when the light intensity seen by the phototransistor circuit was at the midpoint of its range of sensitivity. That is, the circuit produces an analog voltage between 0 and 15 volts based on the amount of light it sees. The midpoint of its sensitivity is when its output equals 7.5 volts. Two dimensions were measured along the horizontal and vertical directions: the width of the hole and the width of the separa-

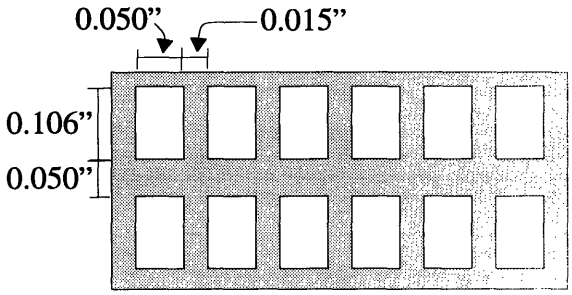


Figure 7.5: Relevant product casing dimensions

tions between adjacent holes, as shown in Figure 7.5. The nominal dimensions of the rectangular hole are 0.036" by 0.106". The gap between adjacent holes along the horizontal axis is 0.015". The gap between the adjacent holes along the vertical axis is 0.050". Five measurements were done for each dimension at each light intensity. Table 7.1 lists the data from the three sets of measurements made.

relative intensity	number	x hole (0.050" nominal)	x separation (0.015" nominal)	z hole (0.106" nominal)	z separation (0.050" nominal)
500 mV	1	0.042"	0.008"	0.076"	0.054"
	2	0.042	0.008	0.077	0.054
	3	0.042	0.008	0.078	0.053
	4	0.042	0.008	0.077	0.054
	5	0.042	0.008	0.077	0.054
560 mV	1	0.044	0.006	0.090	0.047
	2	0.044	0.006	0.090	0.047
	3	0.044	0.006	0.090	0.047
	4	0.044	0.006	0.090	0.047
	5	0.044	0.006	0.090	0.047
580 mV	1	0.046	0.004	0.096	0.046
	2	0.046	0.004	0.096	0.046
	3	0.046	0.004	0.096	0.046
	4	0.046	0.004	0.096	0.046
	5	0.046	0.004	0.096	0.046

Table 7.1: Table of experimental data

The three light intensities were measured in millivolts since this was the output measured by the oscilloscope and was convenient to use. The actual absolute values of the intensities from the lamp were not relevant for these experiments, just the relative differences. The three intensities that were used may seem close together, but in fact, each level produced visibly different amounts of light. For this range of light intensities used, the output would fully swing between 0 and 15 volts when the sensor went from being fully exposed to being fully occluded from the light. Light intensities above and below this range easily saturated the sensor at 0 and 15 volts, respectively. Thus, the experiments were arranged and the lamp intensities were chosen such that the phototransistor circuit would operate within its range of sensitivity.

7.3 Analysis of Results

For each set of data, the measurements are consistent and repeatable. They are consistent with the actual dimensions of the hole, and there is little variation in the measurements for each set.

The dependency on the light intensity of the backlight, however, is evident. For higher intensities of light from the lamp, the output of the phototransistor reaches its midpoint sooner as the hole is moved across its field of view. Consequently, the measured dimensions of the holes grows bigger for higher intensities, while the measured dimensions of the separations decreases. This dependency on the lamp intensity would pose a problem as the light output of a lamp decays with age. It would also pose a problem when this method of sensing was used for different parts with holes of different dimensions. Some means of normalizing the signal and keeping the detector in its operating range is necessary.

This method of using a small probe with a uniform back light has potential. It may be difficult to determine the perimeter of the hole since the analog signal from the photodetector blurs it. However, it would be fairly simple to determine the center of the hole since opposite edges of the hole are equally "blurred" and taking the average of the two readings would theoretically give the midpoint of the hole.

8 Closing Remarks

Flexible assembly systems provide the versatility to perform many functions but come with added complexity and problems with respect to process control. These problems were addressed in the context of a machine that performs bends. First, the system characteristics were identified and used to create a mathematical model. From this model, it was clear what characteristics needed to be adjusted to achieve the desired "step" response. A discrete-time controller complemented with input shaping was devised that optimized the system for speed and accuracy and accounted for the restrictions imposed by the controller card. The controller and the input shaping techniques were incorporated in the implementation of two algorithms for automated bending.

A secure clamping mechanism was also designed and tested. The electromagnetic clamp, incorporated into the design of the kinematic coupling, would allow the bender module to tightly dock with another module.

The development of the controller and bending algorithm for the bender in its current configuration has reached a limit. Other areas of development can be found in the transport module and the optical sensing for fine part location. The development of the transport module would enable the bender to handle a larger family of products. The development of the optical sensing method discussed in chapter 7 would provide one solution to the difficult task of accurate product location. If developed and integrated with the bender, it would provide the bender with the capability to check for correct part location prior to bending. The method is promising and is attractive because of its low cost and simple and small design. However, many of the complexities of the sensing problem have yet to be resolved.

This thesis was intended to provide clear and fairly complete derivations of the methods that were used since they can be more generally applied to understand and control other systems. For example, the frequency response of many electromechanical systems can be determined as described in chapter 2. Knowing the frequency response provides the ability to see what can be done to resolve control problems and improve performance.

Today, many systems are controlled by digital computers. The design of a discrete controller must properly account for the fact that the system exists in an analog world. Typically the computer converts its digital outputs to an analog one through a zero-order hold (ZOH) interpolator. Chapter 3.1 discussed how to model a system driven by a computer through a ZOH interpolator. For a compensator originally designed continuous time, chapter 4 described how to translate it into discrete-time.

Finally, the design of the electromagnet in chapter 6 was meant to be complete and general enough to be useful for designing electromagnets for other applications. It has many advantages over a mechanical clamp since it requires no special mating surfaces, has no moving parts, and can be easily designed around components. Thus, its design would be simple and compact.

ERROR: syntaxerror
OFFENDING COMMAND: , McGraw-Hill, Inc., N606.75 D 368.38 593.38 360.25 602.75 396 581.75 D N 237 609 257 610

STACK:[1] Astrom, K., Wittenmark, B., *Computer-Controlled Systems: Theory and Design*, Prentice-Hall, Englewood Cliffs, NJ, (1984).

[2] Barrientos, M., "Flexible Automation for New Product Introduction," MIT M.S. Thesis (1995).

[3] Bozorth, R., *Ferromagnetism,* IEEE Press, Piscathaway, NJ (1978)

[4] Crandall, S. H., Dahl, N. C., Lardner, T.J., *An Introduction to the Mechanics of Solids* (2^{nd} edition.), McGraw-Hill, New York, NY (1978).

[5] Dempsey, W., "The Design and Implementation of a Serial Insertion Module for Flexible Assembly," MIT M.S. Thesis (1995).

[6] Forsythe, W., Goodall, R. M., *Digital Control: Fundamentals, Theory, and Practice*